

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



Document AN3
Appl. No. 09/482,682

(11)

EP 0 892 047 A2

(12)

EUROPÄISCHE PATENTANMELDUNG

(43) Veröffentlichungstag:
20.01.1999 Patentblatt 1999/03

(51) Int. Cl.⁶: **C12N 15/12**, C07K 14/47,
A61K 48/00, G01N 33/50

(21) Anmeldenummer: 98112470.4

(22) Anmeldetag: 06.07.1998

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: 09.07.1997 DE 19729211
11.02.1998 DE 19805371

(71) Anmelder:
Hoechst Marion Roussel Deutschland GmbH
65929 Frankfurt am Main (DE)

(72) Erfinder:
• Fleckenstein, Bernhard, Prof. Dr.
91369 Wiesenthau (DE)
• Ensser, Armin, Dr.
90419 Nürnberg (DE)

(54) **Humanes und murines Semaphorin L**

(57) Humanes Semaphorin L (H-SemaL) und korrespondierende Semaphorine in anderen Spezies.

Gegenstand der Erfindung sind neuen Semaphorine, die sich durch eine besondere Domänenstruktur auszeichnen und deren Derivate, Nukleinsäuren (DNA, RNA, cDNA), die für diese Semaphorine kodieren und deren Derivate sowie die Verwendung derselben.

Gegenstand der vorliegenden Erfindung sind Semaphorine mit einer neuen, bisher nicht bekannten und nicht zu erwartenden Domänenstruktur, denen eine biochemische Funktion im Immunsystem zukommt (immunmodulierende Semaphorine). Die erfindungsgemäßen Semaphorine werden als Semaphorine vom Typ L (SemaL) bezeichnet. Sie enthalten ein N-terminales Signalpeptid, eine charakteristische Sema-Domäne und im C-terminalen Bereich des Proteins eine Immunglobulin-ähnliche Domäne und eine hydrophobe Domäne, die eine potentielle Transmembrandomäne darstellt.

EP 0 892 047 A2

Beschreibung

Humanes Semaphorin L (H-SemaL) und korrespondierende Semaphorine in anderen Spezies.

Gegenstand der Erfindung sind neue Semaphorine, die sich durch eine besondere Domänenstruktur auszeichnen und deren Derivate, Nukleinsäuren (DNA, RNA, cDNA), die für diese Semaphorine kodieren und deren Derivate sowie die Herstellung und Verwendung derselben.

Semaphorine wurden erstmals von Koldokin {Kolodkin et al. (1993) Cell 75:1389-1399} als Mitglieder einer konservierten Genfamilie beschrieben.

Inzwischen wurden die Gene bzw. Teile der Gene weiterer Semaphorine kloniert und teilweise charakterisiert. Bisher waren insgesamt 5 humane (H-Sema III, H-Sema V, H-Sema IV, H-SemaB und H-SemaE) {Koldokin et al. (1993); Roche et al. (1996) Onkogene 12:1289-1297; Sekido et al. (1996) Proc. Natl. Acad. Sci. USA 93:4120-4125; Xiang et al. (1996) Genomics 32:39-48; Hall et al. (1996) Proc. Natl. Acad. Sci. USA 39:11780-11785; Yamada et al. (1997) (GenBank Zuordnungs-Nr. AB000220)}, 8 murine (Gene der Maus; M-Sema A bis M-SemaH) {Püschel et al. (1995) Neuron 14:941-948; Messerschmidt et al. (1995) Neuron 14:949-959; Inigaki et al. (1995) FEBS Letters 370:269-272; Adams et al. (1996) Mech. Dev. 57:33-45; Christensen et al. (1996) (Genbank Zuordnungs-Nr. Z80941, Z93948)}, 5 galline (Huhn) (Collapsin-1 bis -5) {Luo et al. (1993); Luo et al. (1995) Neuron 14:1131-1140}, und Gene von Ratte (R-Sema III) {Giger et al. (1996) J. Comp. Neurol. 375:378-392}, Zebrafisch, Insekten (Fruchtfliege (*Drosophila melanogaster*: D-Sema I und D-Sema II), Käfer (*Tribolium confusum*: T-Sema I), Grasshüpfer (*Schistocerca americana*: G-Sema I)) {Kolodkin et al. (1993)}, und Nematoden (*C.elegans*: Ce-Sema) {Roy et al. (1994) (GenBank Zuordnungs-Nr. U15667)} bekannt. Weiterhin besitzen zwei Poxviren (*Vaccinia* (ORF-A39) und *Variola* (ORFA39-homolog)) {Kolodkin et al. (1993)} sowie der Alcelaphine Herpesvirus Typ 1 (AHV-1) (AHV-Sema) {Ensser und Fieckenstein (1995) Gen. Virol. 76:1063-1067} zu Semaphorinen homologe Gene.

Einen Überblick über die bisher identifizierten Semaphorine in verschiedenen Spezies gibt Tabelle 1. In Tabelle 1 sind die Namen der Semaphorine (Spalte 1), die verwendeten Synonyme (Spalte 2), die Spezies aus der das jeweilige Semaphorin isoliert wurde (Spalte 3) sowie, soweit bekannt, Daten zur Domänenstruktur des kodierten Proteins und zur chromosomalen Lokalisation (Spalte 4 in Tabelle 1), die Zuordnungsnummer unter der die Sequenz des Gens in Gendatenbanken, z.B. in einer EST (expressed sequence tags) Datenbank, EMBL (European Molecular Biology Laboratory, Heidelberg) oder NCBI (National Center for Biotechnology Information, Maryland, USA) gespeichert ist und die entsprechende Referenz unter der diese Daten publiziert wurden (Spalte 5 in Tabelle 1), angegeben.

Alle Genprodukte (kodierte Semaphorine) der bisher bekannten Semaphorin Gene weisen ein N-terminales Signalpeptid auf, an dessen C-terminalem Ende sich eine charakteristische Sema-Domäne mit einer Länge von etwa 450 bis 500 Aminosäuren befindet. Innerhalb der Sema-Domäne finden sich stark konservierte Aminosäuremotive und eine Anzahl hochkonservierter Cysteinreste. Die Genprodukte (Semaphorine) unterscheiden sich in den auf die Sema-Domäne folgenden C-terminalen Sequenzen, die aus einer oder mehreren Domänen aufgebaut sind. Sie weisen beispielsweise in diesen C-terminalen Aminosäuresequenzen Transmembrandomänen (TM), Immunoglobulin-ähnliche Domänen (Ig) (konstanter Teil des Immunoglobulins), zytoplasmatische Sequenzen (CP), Prozessierungssignale (P) (beispielsweise mit der Konsensussequenz (RXR), wobei R für die Aminosäure Arginin und X für eine beliebige Aminosäure steht) und/oder hydrophile C-Termini (HPC) auf. Auf der Basis der unterschiedlichen Domänenstruktur im C-Terminus lassen sich die bisher bekannten Semaphorine in 5 verschiedene Untergruppen einteilen (I bis V):

- | | | |
|-----|--------------|---|
| I | | Sezerniert, ohne weitere Domäne (z.B. ORF-A49) |
| II | Ig | Sezerniert (ohne Transmembrandomäne) (z.B. AHV-Sema) |
| III | Ig, TM, CP | Membranverankert mit zytoplasmatischer Sequenz (z.B. CD100) |
| IV | Ig, (P), HPC | Sezerniert mit hydrophilem C-Terminus (z.B. H-Sema III, M-SemaD, Collapsin-1) |
| V | Ig, TM, CP | Membranverankert mit C-terminalem 7 Thrombospondin-Motiv (z.B. M-SemaF und G) |

Ein Rezeptor oder extrazellulärer Ligand für Semaphorine wurde bisher nicht beschrieben. Im Zusammenhang mit Semaphorin-vermittelten Effekten wurden intrazelluläre, heterotrimer GTP-bindende Proteinkomplexe beschrieben. Als ein Bestandteil dieser Proteinkomplexe wurden bei Hühnern sogenannte CRMP-Proteine (Collapsin response mediator protein) identifiziert, welche vermutlich Bestandteil der Semaphorin-induzierten intrazellulären Signalkaskade sind (Goshima et al. (1995) Nature 376: 509-514). Das CRMP62 beispielsweise besitzt Homologie zu unc-33, einem für das gerichtete Axonwachstum essentiellen Nematoden-Protein. Ein humanes Protein mit 98% Aminosäure-Identität zu CRMP62 ist ebenfalls bekannt (Hamajima et al. (1996) Gene 180: 157-163). In Ratten wurden ebenfalls mehrere CRMP-verwandte Gene beschrieben (Wang et al. (1996) Neurosci. 16: 6197-6207).

Die sezernierten oder transmembranen Semaphorine vermitteln repulsive Signale für wachsende Nervenknospen. Sie spielen eine Rolle bei der Entwicklung des zentralen Nervensystems (ZNS) und werden vor allem in Muskel- und Nervengewebe exprimiert (Koldokin et al. (1993); Luo et al. (1993) Cell 75:217-227).

Außer im ZNS konnte eine deutliche Expression von M-SemaG auch auf Zellen des lymphatischen und hämato-

poetischen Systems beobachtet werden, im Gegensatz zum nahe verwandten M-SemaF {Furuyima et al. (1996) J. Biol. Chem. 271: 33376-33381}.

Kürzlich wurden zwei weitere humane Semaphorine identifiziert, H-Sema IV und H-Sema V und zwar in einer Region auf Chromosom 3p21.3, deren Deletion mit verschiedenen Formen von Bronchialkarzinomen assoziiert ist H-Sema IV {Roche et al. (1996), Xiang et al. (1996), Sekido et al. (1996)} ist auf Aminosäureebene etwa zu 50% identisch mit M-SemaE, während H-Sema V {Sekido et al. (1996)} das direkte Homolog zu M-Sema-A ist (86% Aminosäureidentität). Da diese Gene (H-Sema IV und V) im Rahmen von DNA-Sequenzierungs-Projekten der deletierten 3p21.3 Loci gefunden wurden, ist die komplexe Intron-Exon-Struktur dieser beiden Gene bekannt. Die beiden Gene werden in verschiedenen neuronalen und nicht neuronalen Geweben exprimiert.

Ebenfalls erst vor kurzem wurde das zelluläre Oberflächenmolekül CD100 (human), exprimiert und induziert auf aktivierten T-Zellen, als Semaphorin identifiziert (ebenfalls in Tabelle 1 aufgeführt). Es unterstützt die Interaktion mit B-Zellen über den Rezeptor CD40 und den entsprechenden Liganden CD40L. CD100 ist ein membranverankertes Glykoprotein-Dimer von 150kd (Kilodalton). Es wurde eine Assoziation des intrazytoplasmatischen C-Terminus von CD100 mit einer noch unbekannten Kinase beschrieben {Hall et al. (1996)}. Damit ist CD100 das erste und bisher einzige Semaphorin, dessen Expression in Zellen des Immunsystems nachgewiesen werden konnte.

Unter der Fragestellung "Transformierende Gene von Rhadinoviren" wurde das komplette Genom des Alcelaphinen Herpesvirus Typ 1 (AHV-1) kloniert und sequenziert {Ensser et al. (1995)}. AHV-1 ist der Erreger des bösartigen Katarrhalebers, einer mit einem lymphoproliferativen Syndrom einhergehenden, meist fatalen Erkrankung verschiedener Wiederkäuer. Bei der Analyse wurde an einem Ende des viralen Genoms ein offener Leserahmen mit entfernter, aber signifikanter Homologie zu einem Gen von Vacciniavirus (ORF-A39 entspricht VAC-A39 in Ensser et al. (1995)) J. Gen. Virol. 76:1063-1067), welches der Genfamilie der Semaphorine zugerechnet wurde, gefunden. Während das AHV-1 Semaphorin (AHV-Sema) eine gut konservierte Semaphorin-Struktur besitzt, sind die Poxvirus-Gene (ORF-A39 und ORF-A39-homolog, siehe Tabelle 1) C-terminal verkürzt, d.h. bei ihnen ist die konservierte Sema-Domäne nur unvollständig vorhanden.

Ein Datenbankvergleich des gefundenen AHV-Sema mit dbEST (EST (expressed sequence tags)-Datenbank (db)) lieferte jeweils 2 EST-Sequenzen von 2 unabhängigen cDNA-Klonen aus humaner Plazenta (Zuordnungsnummern H02902, H03806 (Klon 151129), Zuordnungsnummern R33439 und R33537 (Klon 135941)). Diese wiesen deutlich höhere Homologie zum AHV-1 Semaphorin auf, als zu den bis dahin beschriebenen neuronalen Semaphorinen.

Gegenstand der vorliegenden Erfindung sind Semaphorine mit einer neuen, bisher nicht bekannten und nicht zu erwartenden Domänenstruktur, denen eine biochemische Funktion im Immunsystem zukommt (immunmodulierende Semaphorine). Die erfindungsgemäßen Semaphorine werden als Semaphorine vom Typ L (SemaL) bezeichnet. Sie enthalten ein N-terminales Signalpeptid, eine charakteristische Sema-Domäne und im C-terminalen Bereich des Proteins eine Immunglobulinähnliche Domäne und eine hydrophobe Domäne, die eine potentielle Transmembrandomäne darstellt.

Die Aminosäure-Sequenz des Signalpeptids kann weniger als 70, vorzugsweise weniger als 60 Aminosäuren und mehr als 20, vorzugsweise mehr als 30 Aminosäuren aufweisen, besonders bevorzugt ist eine Länge von etwa 40 bis 50 Aminosäuren. In einer speziellen Ausführungsform der Erfindung hat das Signalpeptid eine Länge von 44 Aminosäuren, d.h. zwischen den Aminosäuren 44 und 45 befindet sich eine Spaltstelle für eine Signalpeptidase.

Die Sema-Domäne kann eine Länge von 300 bis 700 oder mehr, vorzugsweise von etwa 400 bis 600 Aminosäuren aufweisen. Bevorzugt sind Sema-Domänen mit einer Länge von 450 bis 550 Aminosäuren, vorzugsweise von etwa 500 Aminosäuren. In einer besonderen Ausführungsform der Erfindung schließt sich die Sema-Domäne an das Signalpeptid an, wobei sich die Sema-Domäne vorzugsweise bis zur Aminosäure 545 erstreckt.

Die Immunglobulin-ähnliche Domäne kann eine Länge von etwa 30 bis 110 oder mehr Aminosäuren aufweisen, bevorzugt sind Längen zwischen 50 und 90, besonders bevorzugt etwa 70 Aminosäuren.

Die Transmembrandomäne kann eine Länge von etwa 10 bis 35, vorzugsweise von etwa 15 bis 30, besonders bevorzugt von etwa 20 bis 25 Aminosäuren aufweisen.

Gegenstand der Erfindung sind Semaphorine vom Typ L aus verschiedenen Spezies, insbesondere aus Wirbeltieren, beispielsweise aus Vögeln und/oder Fischen, vorzugsweise aus Säugetieren, beispielsweise aus Primaten, Ratte, Kaninchen, Hund, Katze, Schaf, Ziege, Kuh, Pferd, Schwein, besonders bevorzugt aus Mensch und Maus. Gegenstand der Erfindung sind auch entsprechende Semaphorine aus Mikroorganismen, insbesondere aus pathogenen Mikroorganismen, beispielsweise aus Bakterien, Hefen und/oder Viren, z.B. aus Retroviren, insbesondere aus humanpathogenen Mikroorganismen.

Eine Ausführungsform der Erfindung ist ein entsprechendes humanes Semaphorin (H-SemaL), das ein Signalpeptid, eine Sema-Domäne, eine Immunglobulin-ähnliche Domäne und eine Transmembrandomäne aufweist. Eine spezielle Ausführungsform ist das Semaphorin, das durch die Aminosäuresequenz gemäß Tabelle 4 gegeben ist.

Eine weitere Ausführungsform der Erfindung sind korrespondierende Semaphorine in anderen Spezies, die im Bereich der Sema-Domäne eine Aminosäureidentität größer als 40 %, vorzugsweise größer 50 %, besonders bevorzugt größer 60 % im Bezug auf die Sema-Domäne von H-SemaL (Aminosäuren 45 bis 545 der Sequenz in Tabelle 4)

aufweisen. Aus näher verwandten Spezies (z.B. Primaten, Maus) können die korrespondierenden Semaphorine durch aus Aminosäureidentitäten größer als 70%, vorzugsweise größer als 80 %, besonders bevorzugt größer als 90 % aufweisen. Prozentuale Homologien können beispielsweise mit dem Programm GAP (GCG Programm-Paket, Genetic Computer Group (1991)) bestimmt bzw. berechnet werden.

5 Eine derartige Ausführungsform der Erfindung ist ein korrespondierendes Semaphorin der Maus (murines Semaphorin (M-SemaL)). Beispielsweise enthält dieses die partielle Aminosäuresequenz gemäß Tabelle 5 (murines Semaphorin (M-SemaL)).

Die Erfindung betrifft auch korrespondierende Semaphorine, die eine Aminosäureidentität (über die Gesamtlänge der Aminosäuresequenz des Protein betrachtet) von nur etwa 15 bis 20% bei wenig verwandten Spezies (phylogene-

10 tisch weit voneinander entfernt), vorzugsweise 25 bis 30%, besonders bevorzugt 35 bis 40 % oder eine höhere Identität im Bezug auf die gesamte Aminosäuresequenz von H-SemaL gemäß Tabelle 4 aufweisen.

Die Gene, die für Semaphorine von Typ L kodieren, weisen eine komplexe Exon-Intron-Struktur auf. Diese Gene können beispielsweise zwischen 10 und 20 Exons, vorzugsweise etwa 11 bis 18, besonders bevorzugt 12 bis 16 Exons und eine entsprechende Anzahl von Introns aufweisen. Sie können aber auch die gleiche Anzahl Exons und Introns

15 aufweisen wie das Gen von H-SemaL (13 oder 15 Exons, vorzugsweise 14 Exons). Eine besondere Ausführungsform der Erfindung betrifft das Gen von H-SemaL. Dieses Gen hat vorzugsweise eine Länge von 8888 bis 10000 oder mehr Nukleotiden. Das humane Semaphorin-Gen enthält vorzugsweise die Nukleotid-Sequenz, die in Tabelle 14 gegeben ist oder die Nukleotidsequenz, die in der Datenbank GenBank® unter der Zugangsnummer AF030697 hinterlegt wurde. Diese Nukleotidsequenzen enthalten mindestens 13 Introns. Darüber hinaus weist das humane Semaphorin-Gen am

20 5'-Ende einen zusätzlichen Sequenzbereich auf. Dieser Bereich enthält gegebenenfalls weitere kodierende und nichtkodierende Sequenzen, z.B. ein oder zwei weitere introns bzw. Exons..

Die Versuche zur chromosomalen Lokalisation des humanen Semaphorins vom Typ L ergaben, daß das entsprechende Gen an Position 15q22.3-23 lokalisiert ist Entsprechend wurde das Gen für M-SemaL an Position 9A3.3-B lokalisiert.

25 Als Folge der komplexen Intron-Exon Struktur kann das Primärtranskript der Semaphorin mRNA unterschiedlich gespliced werden, wodurch unterschiedliche Splicevarianten der Semaphorine entstehen. Die aus diesen Splicevarianten translatierten Proteine sind Derivate der erfindungsgemäßen Semaphorine. Sie entsprechen in ihrer Aminosäuresequenz und auch weitgehend in ihrer Domänenstruktur den beschriebenen, erfindungsgemäßen Semaphorinen vom Typ L, sind jedoch gegebenenfalls gegenüber diesen verkürzt. Beispielsweise können Splicevarianten, denen die Transmembrandomäne ganz oder teilweise fehlt, gebildet werden. Ein Semaphorin-Derivat, welches keine oder keine

30 vollständige Transmembrandomäne, aber ein Signalpeptid enthält, kann sezerniert werden und auf diese Weise außerhalb der Zelle lokal oder auch über größere Entfernungen wirken, beispielsweise auf andere Zellen. Eine andere Splicevariante kann beispielsweise keine Sequenz mehr enthalten, die für ein Signalpeptid kodiert und gegebenenfalls auch keine Sequenz, die für eine für hydrophobe Aminosäuresequenz kodiert, die eine potentielle Transmembrandomäne

35 darstellt. Eine Folge wäre, daß dieses Semaphorin-Derivat weder in die Membran eingebaut, noch sezerniert wird (es sei denn über sekretorische Vesikel). Ein solches Semaphorin-Derivat kann an intrazellulären Prozessen, beispielsweise an Signaltransduktionsprozessen beteiligt sein. Auf diese Weise können mit dem gleichen Grundmoleküle (Semaphorine vom Typ L) und den davon abgeleiteten Derivaten (beispielsweise Splicevarianten) vielfältige intra- und extrazelluläre Prozesse reguliert und/oder aufeinander abgestimmt werden.

40 Eine besondere Ausführungsform der Erfindung betrifft Semaphorin-Derivate, die sich von den erfindungsgemäßen Semaphorinen vom Typ L ableiten, die aber keine oder keine vollständige Transmembrandomäne enthalten.

Eine weitere Ausführungsform der Erfindung betrifft Semaphorin-Derivate, die sich von den erfindungsgemäßen Semaphorinen vom Typ L ableiten, die aber kein Signalpeptid enthalten.

Das Signalpeptid kann auch posttranslational abgespalten werden. Dadurch wird ein membranständiges (mit TM-

45 Domäne) oder ein sezerniertes (Splicevariante ohne TM-Domäne) Semaphorin-Derivat mit verkürzter Domänenstruktur gebildet. Ein auf diese Weise posttranslational prozessiertes Semaphorin-Derivat enthält nur noch Sema-Domäne, Ig-Domäne und gegebenenfalls Transmembrandomäne. Eine Signalpeptidschnittstelle kann beispielsweise direkt am Ende des Signalpeptids liegen, sie kann z.B. 40 bis 50 Aminosäuren oder weiter vom Aminoterminus entfernt lokalisiert sein.

50 Ein "verkürztes" (d.h. weniger Domänen enthaltendes) Semaphorin L-Derivat ist von anderen Semaphorinen, die sich nicht von den Semaphorinen vom Typ L ableiten, dadurch zu unterscheiden, daß es eine sehr große (> 90 %) Aminosäureidentität oder eine identische Aminosäuresequenz mit den Semaphorinen vom Typ L in den vorhandenen Domänen aufweist.

Die erfindungsgemäßen Semaphorine können auch in anderer Weise posttranslational modifiziert sein. Beispiels-

55 weise können sie ein-, zwei-, drei-, vier-, fünf-, sechs-, sieben-, acht-, neun- zehn- oder mehrfach glykosyliert (N- und/oder O-glykosyliert) vorliegen. Die Aminosäuresequenzen der Semaphorine können dann ebenso viele oder mehr Konsensussequenzen für potentielle Glykosylierungsstellen aufweisen, vorzugsweise fünf derartige Stellen. Eine Ausführungsform der Erfindung betrifft Semaphorine, bei denen die Glykosylierungsstellen an Positionen lokalisiert sind,

die den Positionen 105, 157, 258, 330 und 602 der H-SemaL Aminosäuresequenz entsprechen (Tabelle 4).

Darüber hinaus können die Semaphorine in Form ihrer phosphorylierten Derivate vorliegen. Semaphorine können die Substrate unterschiedlicher Kinasen sein, beispielsweise können die Aminosäuresequenzen Konsensussequenzen für Protein Kinase C, Tyrosin Kinase und/oder Kreatin Kinasen aufweisen. Weiterhin können die Aminosäuresequenzen der Semaphorine Konsensussequenzen für potentielle Myristylierungsstellen aufweisen. An diesen Stellen können entsprechende Semaphorin-Derivate mit Myristinsäure verestert sein.

Die erfindungsgemäßen Semaphorine vom Typ L und deren Derivate können in Form von Monomeren, Dimeren und/oder Multimeren vorliegen, beispielsweise können zwei oder mehr Semaphorine bzw. deren Derivate über intermolekulare Disulfidbrücken miteinander verbunden sein. Darüber hinaus können sich intramolekulare Disulfidbrücken ausbilden.

Derivate der erfindungsgemäßen Semaphorine sind weiterhin Fusionsproteine. Ein solches Fusionsprotein enthält einerseits ein Semaphorin vom Typ L oder Teile desselben und darüber hinaus ein weiteres Peptid oder Protein bzw. ein Teil desselben. Peptide oder Proteine bzw. Teile derselben können z.B. Epitope-Tags (z.B. His-Tag (6xHistidin), Myc-Tag, flu-Tag), die z.B. zur Aufreinigung der Fusionsproteine verwendet werden können, oder solche, die zur Markierung der Fusionsproteine verwendet werden können, z.B. GFP (green fluorescent protein), sein. Beispiele für Derivate der Semaphorine vom Typ L sind z. B. durch die in den Beispielen beschriebenen Konstrukte gegeben. Die Sequenzen dieser Konstrukte können aus den Tabellen 7 bis 15 gegebenenfalls unter Berücksichtigung der Annotationen zu den Plasmiden entnommen werden.

Weiterer Gegenstand der Erfindung sind Nukleinsäure-Sequenzen, vorzugsweise DNA- und RNA-Sequenzen, die für die erfindungsgemäßen Semaphorine vom Typ L und/oder deren Derivate kodieren, beispielsweise die entsprechenden Gene, die unterschiedlichen Splicevarianten der mRNA, die dazu korrespondierenden cDNAs sowie deren Derivate, z.B. Salze der DNA bzw. RNA Derivate im Sinne der Erfindungen sind Sequenzen oder Teile davon, die z.B. mit molekularbiologischen Methoden verändert und an die jeweiligen Anforderungen angepaßt werden, beispielsweise verkürzte Gene oder Teile der Gene (z. B. Promotorsequenzen, Terminatorsequenzen), cDNAs oder Chimäre derselben, Konstrukte für Expressionen und Klonierungen und deren Salze.

Eine Ausführungsform betrifft die genomischen Sequenzen (Gene) der Semaphorine vom Typ L. Die Erfindung betrifft die Intron- und Exon-Sequenzen und genregulatorische Sequenzen, beispielsweise Promotor-, Enhancer- und Silencer-Sequenzen.

Ein Gegenstand dieser Ausführungsform ist das Gen von H-SemaL bzw. dessen Derivate. Ein Gegenstand der Erfindung ist ein Gen, das die Nukleotidsequenz, die in Tabelle 14 gegeben ist, enthält. Ein weiterer Gegenstand der Erfindung ist das Gen, das die Nukleotidsequenz, die in der Datenbank GenBank® unter der Zugangsnummer AF030697 hinterlegt ist, enthält.

Ein weiterer Gegenstand dieser Ausführungsform ist das Gen von M-SemaL bzw. dessen Derivate.

Ein weiterer Gegenstand der Erfindung ist die cDNA von H-SemaL oder deren Derivate (z. B. Teile der cDNA). Eine besondere Ausführungsform ist die cDNA von H-SemaL gemäß der Nukleotidsequenz in Tabelle 2. Ein weiterer Gegenstand der Erfindung ist die cDNA von H-SemaL, die in der Datenbank GenBank® unter der Zugangsnummer AF030698 hinterlegt ist. Gegenstand der Erfindung sind auch die zu diesen cDNAs korrespondierenden mRNAs bzw. Teile derselben.

Ein weiterer Gegenstand der Erfindung ist die cDNA von M-SemaL oder deren Derivate (z. B. Teile der cDNA). Eine besondere Ausführungsform ist die partielle cDNA-Sequenz von M-SemaL gemäß Tabelle 3 sowie cDNA-Sequenzen, die diese partielle cDNA-Sequenz enthalten. Eine weitere Ausführungsform der Erfindung betrifft die cDNA von M-SemaL, die in der Datenbank GenBank unter der Zugangsnummer AF030699 hinterlegt ist. Gegenstand der Erfindung sind auch die zu diesen cDNAs korrespondierenden mRNAs bzw. Teile derselben.

Die Erfindung beinhaltet auch Allele und/oder individuelle Ausprägungsformen der Gene/mRNAs/cDNAs, die sich nur geringfügig von den hier beschriebenen Semaphorin-Sequenzen unterscheiden und für ein identisches oder nur geringfügig verändertes Protein (Abweichung in der Aminosäure-Sequenz kleiner oder gleich 10%) kodieren (weiteres Beispiel für Derivate). Weitere Beispiele für die Derivate sind durch die in den Beispielen angegebenen Konstrukte gegeben. Die Sequenzen dieser Konstrukte sind in den Tabellen 7 bis 14 dargestellt und können unter Berücksichtigung der Annotation für Plasmide interpretiert werden.

Ein weiterer Gegenstand der Erfindung sind Plasmide, die DNA, die für die Semaphorine des Typs L bzw. deren Derivate kodiert, enthalten. Solche Plasmide können beispielsweise Plasmide mit hohen Replikationsraten sein, die für eine Amplifikation der DNA, z.B. in E. coli geeignet sind.

Eine spezielle Ausführungsform sind Expressionsplasmide, mit denen die Semaphorine bzw. Teile davon oder deren Derivate in prokaryoten und/oder eukaryoten Expressions-Systemen exprimiert werden können. Es sind sowohl Expressionsplasmide die konstitutive als auch solche, die induzierbare Promotoren enthalten, geeignet.

Gegenstand der Erfindung sind auch Verfahren zur Herstellung von Nukleinsäuren, die für Semaphorine vom Typ L oder Derivate derselben kodieren.

Beispielsweise können diese Nukleinsäuren, z. B. DNA oder RNA auf chemischem Weg synthetisiert werden. Insbeson-

dere können diese Nukleinsäuren, z. B. die entsprechenden Gene oder cDNAs bzw. Teile derselben mit der PCR unter Verwendung von spezifischen Primern und geeignetem Ausgangsmaterial als Template (z. B. cDNA aus einem geeigneten Gewebe oder genomische DNA) amplifiziert werden.

Ein konkretes Verfahren zur Herstellung von Semaphorin L cDNA bzw. des H SemaL Gens ist in den Beispielen beschrieben.

Die Erfindung betrifft auch Verfahren zur Herstellung der Semaphorine vom Typ L. Beispielsweise kann ein Semaphorin L oder ein Derivat desselben dadurch hergestellt werden, daß eine entsprechende Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L oder ein Derivat desselben kodiert, in einen Expressionsvektor kloniert und mit diesem rekombinanten Vektor eine geeignete Zelle transformiert wird. Es können beispielsweise prokaryote oder eukaryote Zellen verwendet werden. Die Semaphorine vom Typ L oder deren Derivate können gegebenenfalls auch auf chemischem Weg hergestellt werden.

Darüber hinaus können die Semaphorine vom Typ L bzw. deren Derivate als Fusionsproteine exprimiert werden, beispielsweise mit Proteinen oder Peptiden, die einen Nachweis des exprimierten Fusionsproteins erlauben, z.B. als Fusionsprotein mit GFP (green fluorescent protein). Die Semaphorine können auch als Fusionsproteine mit einem, zwei, drei oder mehreren Epitop-Tags, beispielsweise mit Myc- und/oder His-(6xHistidin) und/oder flu-Tags exprimiert werden. Entsprechend können Plasmide verwendet oder hergestellt werden, die DNA-Sequenzen enthalten, die für diese Fusionsproteine kodieren. Beispielsweise können Semaphorin kodierende Sequenzen in Plasmide kloniert werden, die DNA-Sequenzen enthalten, die für GFP und/oder Epitop-Tags, z.B. Myc-Tag, His-Tag, flu-Tag kodieren. Konkrete Beispiele dafür sind durch die Beispiele und die in den Tabellen aufgeführten Sequenzen ggfs. unter Zuhilfenahme der Annotation zu den Plasmiden gegeben.

Ein weiterer Gegenstand der Erfindung sind Antikörper, die spezifisch die Semaphorine vom Typ L, deren Derivate oder Teile davon binden bzw. erkennen. Dies können beispielsweise polyklonale oder monoklonale Antikörper, die z.B. in Maus, Kaninchen, Ziege, Schaf, Huhn usw. hergestellt werden können, sein.

Eine besondere Ausführungsform dieses Gegenstandes der Erfindung sind Antikörper, die gegen die Epitope die den Aminosäuresequenzen von Position 179 bis 378 bzw. 480 bis 666 der H-SemaL Sequenz gemäß Tabelle 4 entsprechen, gerichtet sind. Die Erfindung betrifft auch ein Verfahren zur Herstellung von spezifischen Anti-Semaphorin L Antikörpern, wobei für die Herstellung Antigene verwendet werden, die genannten Epitope enthalten.

Die Erfindung betrifft auch Verfahren zur Herstellung der Antikörper, vorzugsweise wird dazu ein Fusionsprotein, bestehend aus einem charakteristischen Semaphorin Epitop und einem Epitop-Tag, welches für die anschließende Aufreinigung des rekombinanten Fusionsproteins verwendet werden kann. Das aufgereinigte Fusionsprotein kann anschließend für die Immunisierung verwendet werden. Zur Herstellung des rekombinanten Fusionsproteins wird ein entsprechender rekombinanter Expressionsvektor hergestellt und mit diesem eine geeignete Zelle transformiert. Aus dieser Zelle kann das rekombinante Fusionsprotein isoliert werden. Die Durchführung kann beispielsweise wie in Beispiel 8 beschrieben erfolgen.

Diese Antikörper können beispielsweise zur Aufreinigung der entsprechenden Semaphorine, z.B. von H-SemaL und dessen Derivaten z.B. über Affinitätsäulen oder zum immunologischen Nachweis der Proteine, z.B. im ELISA, im Western-Blot und/oder in der Immunhistochemie verwendet werden. Die Antikörper können auch zur Analyse der Expression von H-SemaL, z. B. in verschiedenen Zelltypen bzw. Zelllinien verwendet werden.

Die cDNA von H-SemaL hat eine Länge von 2636 Nukleotiden (Tabelle 2). Das Genprodukt der H-SemaL-cDNA hat eine Länge von etwa 666 Aminosäuren (Tabelle 4) und weist die typische Domänenstruktur eines Semaphorins vom Typ L auf. Das Genprodukt weist ein N-terminales Signalpeptid (Aminosäuren 1 bis 44), Sema-Domäne (Aminosäure 45 bis ungefähr Aminosäure 545) und Ig (Immunglobulin)-Domäne (etwa Aminosäuren 550 bis 620) sowie am C-terminalen Ende eine hydrophobe Aminosäuresequenz auf, die eine potentielle Transmembrandomäne darstellt. Diese Domänen-Struktur wurde bisher für Semaphorine noch nie beschrieben. Es handelt sich um ein membranassoziiertes, wahrscheinlich an der Zelloberfläche lokalisiertes Glykoprotein einer neuen Untergruppe. Aufgrund dieser bisher nicht bekannten Domänenstruktur können die Semaphorine nun in VI Untergruppen eingeteilt werden:

| | | |
|-----|--------------|--|
| I | | Sezerniert, ohne weitere Domäne (z.B. ORF-A49) |
| II | Ig | Sezerniert (ohne Transmembrandomäne) (z.B. AHV-Sema) |
| III | Ig, TM, CP | Membranverankert mit zytoplasmatischer Sequenz (z.B. CD100) |
| IV | Ig, (P), HPC | Sezerniert mit hydrophilem C-Terminus (z.B. H-Sema III, M-SemaD, Collapsin-1) |
| V | Ig, TM, CP | Membranverankert mit C-terminalem 7 Thrombospondin-Motiv (z.B. M-Sema-F und G) |
| VI | Ig, | TM Membranverankert (z.B. H-SemaL, M-SemaL) |

Die nichtglykosylierte, nichtprozessierte Form von H-SemaL hat ein errechnetes Molekulargewicht von etwa 74,8 kD (74823 Dalton) (berechnet mit PeptideSort, GCG-Programm-Paket). Der Isoelektrische Punkt berechnet sich zu pH=7.56.

Eine mögliche Signalpeptid-Schnittstelle liegt zwischen den Aminosäuren 44 und 45 (Tabelle 3; berechnet mit SignalP

(http://www.cbs.dtu.dk/services/Signal_P), einem auf neuronalen Netzwerken basierenden Programm zur Analyse von Signalsequenzen (Nielsen H. et al. (1997) Protein Engineering 10:1-6)). Dies ergibt für das prozessierte Protein (ohne Signalpeptid) ein Molekulargewicht (MW) von 70,3 kD (70323 Dalton) und einen Isoelektrischen Punkt von pH=7.01.

Die genomische Struktur ist ebenfalls weitgehend geklärt. Das H-SemaL-Gen weist 13 oder 15 oder mehr Exons, vorzugsweise 14 Exons und 12 oder 14 Introns, vorzugsweise 13 Introns auf. Aufgrund dieser komplexen Exon-Intron-Struktur sind unterschiedliche Splice-Varianten möglich. Die mRNA des transkribierten H-SemaL-Gens findet sich im Northern-Blot vor allem in Placenta, Keimdrüsen, Thymus und Milz. Es wurde keine mRNA in neuronalem Gewebe oder in Muskelgewebe nachgewiesen. Ein Hinweis auf eine spezifisch regulierte Expression in Endothelzellen liegt vor.

Durch alternatives Splicing können auch Formen von H-SemaL mit intrazytoplasmatischen Sequenzen entstehen, die eine Rolle in der intrazellulären Signaltransduktion spielen, ähnlich wie z.B. bei CD100.

Ebenfalls möglich wären durch alternatives Splicing entstehende, sezernierte Formen von H-SemaL, analog zum viralen AHV-Sema.

Nukleotid- und Aminosäuresequenzanalysen wurden mit Hilfe des GCG Programm-Pakets (Genetics Computer Group (1991) Program manual for the GCG package, version 7, 575 Science Drive, Wisconsin, USA 53711), FASTA (Pearson und Lipman (1988) Proc. Natl. Acad. Sci. 85, 2444-2448) und BLAST-Program (Gish und States (1993) Nat. Genet. 3, 266-272; Altschul et al. (1990) J. Mol. Biol. 215, 403-410) durchgeführt. Diese Programme wurden auch für Sequenzvergleiche mit GenBank (Version 102.0) und Swiss Prof (Version 34.0) verwendet.

Posttranslationale Modifikationen wie Glykosylierung und Myristylierung von H-SemaL sind ebenfalls möglich. Konsensus-Sequenzen für N-Glykosylierungsstellen wurden mit Hilfe des Programms Prosite (GCG Programm-Paket) an den Positionen 105, 157, 258, 330, 602 der Aminosäuresequenz von H-SemaL (gemäß Tabelle 4) gefunden, solche für Myristylierung an den Positionen 114, 139, 271, 498, 499, 502, 654 (Konsensus-Sequenz: G-(E, D, R, K, H, P, F, Y, W) x (S, T, A, G, C, N)-(P)). Darüber hinaus enthält die Aminosäuresequenz von H-SemaL mehrere Konsensus-Sequenzen für potentielle Phosphorylierungsstellen durch unterschiedliche Kinasen. Deshalb kann davon ausgegangen werden, daß H-SemaL das Substrat unterschiedlicher Kinasen sein kann, z.B. Phosphorylierungsstellen für Kreatin-Kinase 2, Protein-Kinase C und Tyrosin-Kinase.

Vorausgesagte Kreatin-Kinase 2-Phosphorylierungs-Stellen (Konsensussequenz Ck2: (S,T)x2(D,E)) (Prosite, GCG) an den Positionen 119, 131, 173, 338, 419, 481 der Aminosäuresequenz. Vorausgesagte Protein-Kinase-C-Phosphorylierungs-Stellen (Konsensussequenz PkC: (S,T)x(R,K)) (Prosite, GCG) an den Positionen 107, 115, 190, 296, 350, 431, 524, 576 der Aminosäuresequenz. Vorausgesagte Tyrosin-Kinase-Phosphorylierungs-Stelle (Konsensussequenz: (R,K)x2,3(D,E)x2,3(Y)) (Prosite GCG) an Position 205 der Aminosäuresequenz.

Die Konsensussequenzen sind im Einbuchstabencode für Aminosäuren angegeben.

Ein für Integrine charakteristisches "RGD"-Motiv (Arginin-Glycin-Asparaginsäure) findet sich an Position 267.

Die Glykosylierungsstellen sind gut konserviert zwischen viralem AHV-Sema, H-SemaL und (soweit bekannt) M-SemaL.

Eine Di- oder Multimerisierung von H-SemaL ist möglich und wurde bei anderen Semaphorinen wie CD100 beschrieben (Hall et al. (1996)). Das CD100 Molekül ist ebenfalls ein membranverankertes Glykoprotein-Dimer von 150kd. CD100 ist jedoch nicht nahe verwandt mit dem erfindungsgemäßen humanen Semaphorin (H-SemaL).

Die partielle cDNA-Sequenz von M-SemaL hat eine Länge von 1195 Nukleotiden. Diese Sequenz kodiert für ein Protein mit 394 Aminosäuren. Diese 394 Aminosäuren entsprechen den Aminosäuren 1 bis 396 von H-SemaL. Das Signalpeptid im M-SemaL erstreckt sich über die Aminosäuren 1 bis 44 (genau wie im H-SemaL). Die Sema-Domäne beginnt bei der Aminosäure 45 und erstreckt sich bis zum Ende bzw. wahrscheinlich über das Ende der Sequenz gemäß Tabelle 4 hinaus.

Multiple Alignments wurden mit Hilfe des Programms Clustal W (Thompson et al. (1994)) durchgeführt. Diese Alignments wurden manuell weiterbearbeitet mit Hilfe von SEAVIEW (Galtier et al. (1996) Comput Appl. Biosci 12, 543-548). Die phylogenetischen Entfernungen wurden mit Clustal W (Thompson et al. (1994)) bestimmt.

Ein Vergleich der Proteinsequenzen der bekannten und der neuen Semaphorine und eine phylogenetische Analyse dieser Sequenzen zeigt, daß sich die Gene entsprechend ihrer phylogenetischen Verwandtschaft einteilen lassen. Hier fließt natürlich die C-terminale Domänenstruktur der entsprechenden Semaphorin-Subtypen als entscheidender Faktor mit ein, weshalb Semaphorine der gleichen Untergruppen in der Regel auch phylogenetisch näher verwandt sind, als Semaphorine unterschiedlicher Untergruppen. Einfluß hat auch, aus welcher Spezies das Semaphorin isoliert wurde, d.h. ob die entsprechenden Spezies phylogenetisch nahe miteinander verwandt sind oder nicht.

Eine phylogenetische Analyse (vergl. Figur 3) der bekannten Semaphorin Aminosäuresequenzen (vollständige Sequenzen und/oder Teilsequenzen, wobei die Aminosäuresequenzen für H-SemaL und M-SemaL gemäß den Tabellen 4 und 5 verwendet wurden, für alle anderen Sequenzen, die unter den Zugangsnummern gespeicherten Sequenzen bzw. die von diesen Sequenzen abgeleiteten Aminosäuresequenzen)) mittels des Programms CLUSTALW (Thompson J.D. et al. (1994) Nucleic Acids Res. 22:4673-4680) zeigt, daß die Aminosäuresequenzen von H-SemaL und M-SemaL phylogenetisch nahe miteinander verwandt sind und eine eigene phylogenetische Gruppe bilden. H-SemaL und M-SemaL wiederum sind phylogenetisch am nächsten verwandt mit AHV-Sema und Vac-A39. Sie sind

untereinander deutlich näher verwandt, als mit irgendeinem anderen bisher bekannten Semaphorin. Die Analyse zeigt weiterhin, daß auch andere Semaphorine phylogenetisch nahe miteinander verwandt sind und eigene Gruppen innerhalb der Semaphorine bilden. Beispielsweise fallen die Semaphorine, die sezerniert werden, z.B. H-Sema III, IV, V, und E in eine phylogenetische Gruppe. Zu dieser Subfamilie gehören auch deren Homologe in anderen Spezies, während das humane (transmembrane) CD100 mit dem entsprechenden Maus-Homologen (M-Sema G2) und mit Collapsin-4 in eine phylogenetische Gruppe fällt.

Im Bezug auf die gesamten Aminosäuresequenzen liegen die beobachteten Homologien innerhalb der phylogenetischen Gruppen zwischen etwa 90% und 80% Aminosäureidentität im Bezug auf sehr nahe verwandte Gene wie z.B. H- und M-SemaE oder -III/D und etwas weniger als 40% bei wenig verwandten Genen der Semaphorine. Innerhalb der Sema-Domäne liegt die beobachtete Aminosäureidentität um einige Prozent höher, und durch ihren großen Anteil am Gesamtprotein (50-80% des Proteins gehören zur Sema-Domäne) der Aminosäuresequenz beeinflußt diese wesentlich die Gesamtidentität.

H-SemaL ist, über das Gesamtprotein berechnet, zu 46% identisch mit AHV-Sema, wird dagegen die Sema-Domäne allein betrachtet, dann beträgt die Aminosäureidentität 53%. Dies ist höher als z.B. zwischen den verwandten M-SemaB und -C (37% Identität im Bezug auf das Gesamtprotein, 43% Identität im Bezug auf die Sema-Domäne), ähnlich wie M-SemaA und -E (43% Gesamtprotein, 53% Sema-Domäne). Die Aminosäureidentität zwischen der partiellen M-SemaL Sequenz (Tabelle 6) und H-SemaL (Tabelle 5) liegt im Bereich der Sema-Domäne bei 93%, so daß davon ausgegangen werden kann, daß es sich um das entsprechend homologe Gen der Maus handelt.

Korrespondierende Semaphorine zu H-SemaL und M-SemaL in anderen Spezies können innerhalb der Sema-Domäne eine Aminosäureidentität größer als 40% im Bezug auf H-SemaL aufweisen. Bei den nahe verwandten Wirbeltieren (Säuger, Vögel) können sogar Aminosäureidentitäten über 70% angetroffen werden.

Es handelt sich um eine neue Subfamilie von Semaphorinen mit größerer Aminosäureidentität zu dem viralen AHV-Sema als zu den bisher bekannten humanen bzw. murinen Semaphorinen, und mit einer für humane Semaphorine bisher nicht bekannten C-terminalen Struktur. Diese neuen Semaphorine (Mitglieder der Subfamilie) zeichnen sich dadurch aus, daß sie aufgrund ihrer Domänenstruktur in die Untergruppe IV fallen und/oder die gleiche phylogenetische Gruppe fallen wie H-SemaL und M-SemaL und/oder im Bezug auf die gesamte Aminosäuresequenz zu H-SemaL eine Aminosäureidentität von mindestens 30 bis 40 %, vorzugsweise 50 bis 60 %, besonders bevorzugt 70 bis 80 % oder eine größere Identität aufweisen und/oder im Bezug auf die Sema-Domäne eine Aminosäureidentität mit H-SemaL von mindestens 70 %, vorzugsweise größer 80 %, besonders bevorzugt größer 90 % aufweisen.

Den Semaphorinen vom Typ L kommt auch eine andersartige biochemische Funktion zu. Eine neue Funktion dieser Semaphorine liegt in der Modulation des Immunsystems.

Das nächste Verwandte von H-SemaL ist das virale AHV-Semaphorin (AHV-Sema). Dieses ist von ähnlicher Größe, besitzt aber im Gegensatz zum H-SemaL keine Transmembrandomäne. Das AHV-Sema wird vermutlich von virusinfizierten Zellen sezerniert, um den H-SemaL äquivalenten Rezeptor (Semaphorin von Typ L im Streifengnu) im natürlichen Wirt (Streifengnu) zu blockieren, und so dem Angriff des Immunsystems zu entgehen. Ferner ist eine Funktion als repulsives Agens (Chemorepellent) für Zellen des Immunsystems denkbar.

Die biochemische Funktion der neuen Semaphorine vom Typ L und deren Derivate ist als generell immunmodulierend und/oder entzündungsmodulierend anzusehen. Einerseits können sie

A) als die Immunantwort hemmende Moleküle entweder lokal, z.B. als Transmembranprotein an der Oberfläche von Zellen oder auch über größere Entfernungen, z.B. wenn sie durch Prozessierung (z.B. Proteasen) oder alternatives Splicing sezerniert werden, z.B. durch Diffusion im Gewebe, ihre Wirkung als Chemorepellent und/oder Immunsuppressivum entfalten.

Beispielsweise kann die Expression dieser neuen Semaphorine vom Typ L z.B. an der Oberfläche der Zellen der Gefäßendothelien das Leukozyten-Attachment und deren Migration durch die Gefäßwand verhindern. Den neuen Semaphorinen kann eine Rolle bei der Aufrechterhaltung von Schrankenwirkungen, z.B. zur Verhinderung von Infektionen in besonders "wichtigen" oder exponierten Organen, beispielsweise zur Aufrechterhaltung der Blut-Hirn-Schranke, des Plazentarkreislaufs und/oder anderen immunologisch privilegierten Orten (z.B. Pancreas-Inseln) und/oder beim Schutz vor Autoimmunerkrankungen zukommen. Darüber hinaus können die neuen Semaphorine und/oder ihre Derivate in verschiedenen Geweben auch an repulsiven Signalen, beispielsweise für Zellen des Immunsystems (z.B. Leukozyten) als Schutz gegen versehentliche Aktivierung von Abwehrmechanismen beteiligt sein.

B) Weiterhin können den neuen Semaphorinen und/oder deren Derivaten Funktionen als akzessorische Moleküle zukommen. An der Zelloberfläche exprimiert können sie beispielsweise an der Interaktion mit Zellen des Immunsystems im Rahmen der Aktivierung von Abwehrmechanismen z. B. bei Virusinfektionen beteiligt sein.

Dadurch ergeben sich mehrere Verwendungsmöglichkeiten für die neuen Semaphorine vom Typ L und deren Derivate.

vaten sowie den für diese Proteine kodierenden Nukleinsäuren.

Funktion A): Es handelt sich um ein immunsuppressives und/oder entzündungshemmendes Prinzip: Zahlreiche potentielle Anwendungsmöglichkeiten liegen in den Bereichen Organtransplantation, Entzündungstherapie, Immuntherapie und Gentherapie.

5 Beispielsweise können mit Hilfe der Semaphorin-kodierenden DNA oder Derivaten derselben nichthumane, transgene Tiere hergestellt werden.

Eine Anwendungsmöglichkeit dieser Tiere liegt in der Hemmung der Transplantatabstoßung in transgenen Modellen für Organtransplantationen. Beispielsweise können transgene, gegen Abstoßung geschützte tierische Organe für Xenotransplantationen hergestellt werden. Dies sollte z.B. auch zusammen mit anderen Transgenen (z.B. Komplementregulatoren wie DAF oder CD59)-möglich sein.

10 Eine weitere Anwendung liegt in der Herstellung von nicht-humanen Knock-out Tieren, beispielsweise von Knock-out Mäusen („Laboratory Protocols for Gene-Targeting“, Torres and Kühn (1997) Oxford University Press, ISBN 0-19-963677-X): Durch Knock-out des Mausgens M-SemaL können z.B. weitere Funktionen des Gens aufgefunden werden. Sie stellen auch potentielle Modellsysteme für entzündliche Erkrankungen dar, falls die Mäuse ohne Semaphorin-Gen lebensfähig sind. Sollte M-SemaL für die Immunmodulation wichtig sein, so sind vermehrt solche Mäuse zu erwarten.

15 Weiterhin können nicht-humane Knock-in-Tiere, beispielsweise Mäuse, hergestellt werden. Dabei wird z.B. M-SemaL durch normales/verändertes H-SemaL oder verändertes M-SemaL (z.B. Integration der neuen Semaphorin-Subtypen unter der Kontrolle von konstitutiven und/oder induzierbaren Promotoren) ersetzt. Solche Tiere können z.B. für die Suche nach weiteren Funktionen der neuen Semaphorine, z.B. Funktionen des humanen Gens oder Derivaten dieser Gene dienen oder zur Identifizierung und Charakterisierung von immunmodulierenden Wirkstoffen benutzt werden.

20 Verwendung von z. B. Nukleinsäuren, die für Semaphorine vom Typ L oder Derivate derselben kodieren, zur Herstellung von z.B. rekombinanten Immunsuppressiven, anderen löslichen Proteinen oder Peptiden die sich von der Aminosäuresequenz der Semaphorine vom Typ L, z.B. von H-SemaL oder den entsprechenden Nukleinsäuren, z.B. Genen ableiten. In ähnlicher Weise können auch Agonisten mit struktureller Ähnlichkeit hergestellt werden. Diese immunsuppressiven Wirkstoffe/Agonisten können bei Autoimmunerkrankungen und entzündlichen Erkrankungen und/oder Organtransplantationen eingesetzt werden.

25 Gentherapie mit Semaphorinen vom Typ L, z.B. mit Nukleinsäuren, die für H-SemaL oder deren Derivate kodieren, z. B. mittels viraler oder nichtviraler Methoden. Einsatz bei Autoimmunerkrankungen und entzündlichen Erkrankungen, der Transduktion von Organen sowie vor/während/nach Transplantationen zur Verhinderung der Transplantatabstoßung.

30 Insbesondere können die neuen Semaphorine und/oder die für diese Semaphorine kodierenden Nukleinsäuren und Derivate derselben, insbesondere H-SemaL, für H-SemaL kodierende DNA und Derivate derselben in einem Verfahren zum Screening von Wirkstoffen, insbesondere zur Identifizierung und Charakterisierung von immnunodulierenden Wirkstoffen, eingesetzt werden.

35 Funktion B): H-SemaL ist ein akzessorisches Molekül, das an der Zelloberfläche exprimiert wird und an der Interaktion mit Zellen, z.B. des Immunsystems, z.B. als akzessorisches Molekül in der Aktivierung von Signalwegen, beteiligt ist. Ein virales Gen bzw. das Genprodukt eines viralen oder anderen pathogenen Gens z.B. mikrobiologischen Ursprungs könnte z.B. als kompetitiver Inhibitor dieses akzessorischen Moleküls wirken. Eine Anwendung für die neuen Semaphorine liegt bei dieser Funktion ebenfalls im Bereich der Organtransplantation, Entzündungstherapie, Immuntherapie und/oder Gentherapie.

40 Beispielsweise können die neuen Semaphorine in einem Verfahren zum Screening von antagonistischen Wirkstoffen bzw. Inhibitoren verwendet werden. Auf diese Weise identifizierte Wirkstoffe können dann z.B. zur Blockade des Semaphorin-Rezeptors eingesetzt werden. Lösliche und/oder sezernierte H-SemaL Antagonisten bzw. Inhibitoren können beispielsweise chemische Substanzen oder die neuen Semaphorine bzw. Derivate derselben selbst sein (z.B. Teile/verkürzte Formen derselben, beispielsweise ohne Membrandomäne oder als Ig-Fusionsproteine oder von diesen abgeleitete Peptide, die geeignet sind, den korrespondierenden Rezeptor zu blockieren). Auf diese Weise identifizierte, spezifische Antagonisten und/oder Inhibitoren können beispielsweise kompetitiv wirken und bei der Hemmung der Abstoßung z.B. in transgenen Modellen für Organtransplantationen und bei Autoimmunerkrankungen, entzündlichen Erkrankungen und Organtransplantationen eingesetzt werden. Nukleinsäuren, z. B. DNA, die für die neuen Semaphorine kodieren bzw. deren mit Hilfe von molekularbiologischen Methoden erzeugte Derivate können beispielsweise für die Herstellung nichthumaner, transgener Tiere verwendet werden. Eine Überexpression von H-SemaL kann in diesen transgenen Tieren zu vermehrter Anfälligkeit für Autoimmunerkrankungen und/oder entzündlichen Erkrankungen führen. Solche transgenen Tiere eignen sich damit zum Screening von neuen, spezifischen, immunmodulierenden Wirkstoffen.

55 Solche Nukleinsäuren können ebenso für die Herstellung von nicht-humanen Knock-out Tieren, beispielsweise Knock-out Mäusen, bei denen das Mausgen M-SemaL ausgeschaltet wird, verwendet werden. Solche Knock-out Tiere können für die Suche nach weiteren biochemischen Funktionen des Gens eingesetzt werden. Sie stellen auch potentielle Modellsysteme für entzündliche Erkrankungen dar, falls die Mäuse ohne das M-SemaL Gen lebensfähig sind.

Diese DNA kann ebenso zur Herstellung von nicht-humanen Knock-in Tieren, beispielsweise Mäusen verwendet werden. Dabei wird das M-SemaL-Gen durch ein verändertes M-SemaL Gen/cDNA oder ein gegebenenfalls verändertes, z.B. mutiertes Semaphorin Typ L Gen/cDNA einer anderen Spezies, z.B. von H-SemaL ersetzt. Solche transgenen Tiere können für die Suche nach weiteren Funktionen der erfindungsgemäßen Semaphorine verwendet werden.

Die Erfindung betrifft auch die Verwendung der Semaphorine vom Typ L und deren Derivate, sowie der für diese Proteine kodierenden Nukleinsäuren, z. B. Gene/cDNAs und deren Derivate und/oder mit Hilfe dieser Semaphorine identifizierter Wirkstoffe zur Herstellung von Arzneimitteln. Beispielsweise können Arzneimittel hergestellt werden, die in der Gentherapie angewendet werden können und die Agonisten und/oder Antagonisten der Expression der Semaphorine vom Typ L, beispielsweise von H-SemaL, enthalten. Dazu können z. B. virale und/oder nichtvirale Methoden verwendet werden. Diese Arzneimittel können z.B. bei Autoimmunerkrankungen und entzündliche Erkrankungen, Organtransplantationen vor und/oder während und/oder nach der Transplantation zur Verhinderung der Abstoßung eingesetzt werden.

Die für die neuen Semaphorine kodierenden Nukleinsäuren, z. B. Gene, cDNAs und deren Derivate können auch als Hilfsmittel in der Molekularbiologie eingesetzt werden.

Darüberhinaus können die neuen Semaphorine, insbesondere H-SemaL und Nukleinsäuren, z. B. Gene/cDNAs derselben in Verfahren zum Screening neuer Wirkstoffe eingesetzt werden. Modifizierte Proteine und/oder Peptide, die sich z. B. von H-SemaL und/oder M-SemaL ableiten, können zur Suche nach dem entsprechenden Rezeptor und/oder dessen Antagonisten bzw. Agonist in funktionellen Tests, beispielsweise mittels Expressionskonstrukten von H-SemaL und Homologen eingesetzt werden.

Die Erfindung betrifft auch die Verwendung eines Semaphorins vom Typ L oder einer Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert in einem Verfahren zur Identifizierung von pharmakologischen Wirkstoffen, insbesondere von immunmodulierenden Wirkstoffen.

Die Erfindung betrifft auch Verfahren zur Identifizierung von Wirkstoffen, wobei ein Semaphorin vom Typ L oder ein Derivat derselben bzw. eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert oder ein Derivat derselben eingesetzt wird, um pharmakologische Wirkstoffe, z. B. immunmodulierende Wirkstoffe zu identifizieren. Beispielsweise betrifft die Erfindung ein Verfahren, bei dem ein Semaphorin vom Typ L unter definierten Bedingungen mit einem zu untersuchenden Wirkstoff inkubiert wird und parallel ein zweiter Ansatz ohne den zu untersuchenden Wirkstoff, aber unter ansonsten gleichen Bedingungen durchgeführt wird und dann die inhibierende bzw. aktivierende Wirkung des zu untersuchenden Wirkstoffs bestimmt wird.

Beispielsweise betrifft die Erfindung auch Verfahren zur Identifizierung von Wirkstoffen, wobei eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert oder ein Derivat derselben unter definierten Bedingungen in Gegenwart eines zu untersuchenden Wirkstoffs exprimiert wird und das Ausmaß der Expression bestimmt wird. Gegebenenfalls können auch in einem solchen Verfahren zwei oder mehrere Ansätze parallel unter gleichen Bedingungen durchgeführt werden, wobei die Ansätze aber unterschiedliche Mengen des zu untersuchenden Wirkstoffs enthalten.

Beispielsweise kann der zu untersuchende Wirkstoff die Transkription und/oder die Translation inhibieren oder aktivieren.

Das Semaphorin vom Typ L kann wie seine viralen Homologen an das neu beschriebene Rezeptormolekül VESPR (Comeau et al, (1998) Immunity, Vol. 8, 473-482) binden und kann vermutlich in Monozyten eine Induktion von Zell-Adhäsionsmolekülen wie ICAM-1 und Zytokinen wie Interleukin-6 und Interleukin-8 bewirken. Dies kann zu deren Aktivierung und zur Zellaggregation führen. Das Expressionsmuster des VESPR-Rezeptors zeigt teilweise interessante Parallelen zu H-SemaL, z.B. eine starke Expression in Placenta und eine deutliche Expression in Milzgewebe. Interaktionen mit weiteren, noch unbekannten Rezeptoren der Plexin-Familie, oder anderen Rezeptoren sind möglich. Auch eine Interaktion mit sich selbst oder anderen semaphorinähnlichen Molekülen ist möglich. Eine Interaktion der Semaphorine vom Typ L kann insbesondere über eine konservierte Domäne im C-terminalen Bereich der Sema-Domäne stattfinden.

Zu den Annotation Plasmiden:

pMelBacA-H-SemaL (6622bp) in pMelBacA (Invitrogen, De Schelp, NL)(SEQ ID NO.42). Nukleotid 96-98 ATG - Startkodon, Nukleotid 96-168 Mellitin Signal-Sequenz, Nukleotid 168-173 BamHI Schnittstelle (PCR/Klonierung), Nukleotid 171-1998 Leserahmen SemaL Aminosäuren 42-649 (ohne eigene Signal-Sequenz und ohne Transmembransequenz), Nukleotid 1993-1998 EcoRI Schnittstelle (PCR/Klonierung) und Nukleotid 1992-1994 Stop Codon

Plasmid pCDNA3.1-H-SemaL-MychisA (7475 bp) (SEQ ID NO. 35):

Nukleotid 954-959 BamHI Schnittstelle (Klonierung), Nukleotid 968-970 ATG SemaL, Nukleotid 968-2965 Leserahmen SemaL, Nukleotid 2963-2968 Pml I Schnittstelle, Nukleotid 2969-2974 HindIII Schnittstelle, Nukleotid 2981-3013 Myc-Tag, Nukleotid 3026-3033 6xHis-Tag, Nukleotid 3034-3036 Stop Codon.

Plasmid pCDNA3.1-H-SemaL-EGFP-MyhisA (8192 bp): (SEQ ID NO. 36):

Nukleotid 954-959 BamHI Schnittstelle (Klonierung), Nukleotid 968-970 ATG SemaL, Nukleotid 968-2965 Leserahmen SemaL, Nukleotid 2963-2965 halbe Pml I Schnittstelle, Nukleotid 2966-3682 Leserahmen EGFP (in Pml I Kloniert), Nukleotid 3683-3685 halbe Pml I Schnittstelle, Nukleotid 3685-3691 HindIII, Nukleotid 3698-3730 Myc-Tag, Nukleotid 3743-3760 6xHis-Tag, und Nukleotid 3761-3763 Stop Codon

Plasmid pIND-H-SemaL-EA (7108 bp) in Vektor pIND (Invitrogen, De Schelp, NL) (SEQ ID No. 38):

Nukleotid 533-538 BamHI Schnittstelle (Klonierung), Nukleotid 546-548 ATG SemaL, Nukleotid 546-Leserahmen SEMAL, Nukleotid 2542-2547 Pml I Schnittstelle, Nukleotid 2546-2553 HindIII Schnittstelle und Nukleotid 2563-2565 Stop Codon.

Plasmid pIND-H-SemaL-EE (Gesamtlänge 7102 bp) in Vektor pIND (Invitrogen, De Schelp, NL) (SEQ ID No. 37):

Nukleotid 533-538 BamHI Schnittstelle (Klonierung), Nukleotid 546-548 ATG SemaL, Nukleotid 546-Leserahmen SemaL, Nukleotid 2542-2547 Pml I Schnittstelle, Nukleotid 2548-2553 HindIII Schnittstelle, Nukleotid 2560-2592 Myc-Tag, Nukleotid 2605-2622 6xHis-Tag und Nukleotid 2623-2625 Stop Codon.

Plasmid pQE30-H-SemaL-179-378.seq (4019 bp) in Vektor pQE30 (Qiagen, Hilden) entspricht pQE30-H-SemaLBH (SEQ ID No. 39):

Nukleotid 115-117 ATG, Nukleotid 127-144 6xHis-Tag, Nukleotid 145-750 BamHI-HindIII PCR-Fragment SemaL Aminosäuren (aa) 179-378 und Nukleotid 758-760 Stop Codon.

Plasmid pQE31-H-SemaL- (SH (3999 bp) in Vektor pQE31 (Qiagen, Hilden) (SEQ ID No. 40):

Nukleotid 115-117 ATG, Nukleotid 127-144 6xHis-Tag, Nukleotid 147-152 BamHI, Nukleotid 159-729 SacI-HindIII Fragment SemaL (C-terminal) aa480-666 und Nukleotid 734-736 Stop Codon.

Beispiele:

Versuchsbedingungen, die in den Beispielen Anwendung finden:

Verwendete PCR-Programme:

Taq52-60 (mit Ampli-Taq^R-Polymerase, Perkin Elmer, Weil der Stadt, Deutschland)

| | |
|----------------------------|-----------|
| 96°C/60s | 1 Zyklus |
| 96°C/15s-52°C/20s-70°C/60s | 40 Zyklen |
| 70°C/60s | 1 Zyklus |

Taq60-30

| | |
|----------------------------|-----------|
| 96°C/60s | 1 Zyklus |
| 96°C/15s-60°C/20s-70°C/30s | 35 Zyklen |
| 70°C/60s | 1 Zyklus |

Taq60-60

| | |
|----------------------------|-----------|
| 96°C/60s | 1 Zyklus |
| 96°C/15s-60°C/20s-70°C/60s | 35 Zyklen |
| 70°C/60s | 1 Zyklus |

Taq62-40

| | |
|----------------------------|-----------|
| 96°C/60s | 1 Zyklus |
| 96°C/15s-62°C/20s-70°C/40s | 35 Zyklen |
| 70°C/60s | 1 Zyklus |

Verwendete Reaktionsbedingungen für PCR mit Taq-Polymerase:

50µl Reaktionsansätze mit 100-200ng Template, 200µM dNTP, 0,2-0,4 µM je Primer, 2.5U Ampli-Taq^R, 5µl des mit-

gelieferten 10x Reaktionspuffers

Verwendete Programme für:

- 5 1. XL62-6 (mit Expand-Long Template PCR System^R, Boehringer Mannheim, Deutschland)

| | | |
|----|--|-----------|
| | 94°C/60s | 1 Zyklus |
| | 94°C/15s-62°C/30s-68°C/6min | 10 Zyklen |
| | 94°C/15s-62°C/30s-68°C/(6min+15s/Zyklus) | 25 Zyklen |
| 10 | 68°C/7min | 1 Zyklus |

2. XL62-12 (mit Expand-Long Template PCR System^R, Boehringer Mannheim, Deutschland)

| | | |
|----|---|-----------|
| | 94°C/60s | 1 Zyklus |
| 15 | 94°C/15s-62°C/30s-68°C/12min | 10 Zyklen |
| | 94°C/15s-62°C/30s-68°C/(12min+15s/Zyklus) | 25 Zyklen |
| | 68°C/7min | 1 Zyklus |

Reaktionsbedingungen für PCR mit Expand-Long Template PCR System

20 50µl Reaktionsansätze mit 100-200ng Template, 500µM dNTP, 0,2-0,4 µM je Primer, 0,75µl Enzym-Mix, 5µl des mitgelieferten 10x Reaktionspuffers Nr. 2.

Beispiel 1:

25 Ausgehend von Sequenzen des AHV-Sema (Ensser u. Fleckenstein (1995), J. General Virol. 76: 1063-1067) wurden PCRs und RACE-PCRs durchgeführt. Als Ausgangsmaterial hierfür diente humane cDNA aus Plazenta-Gewebe, an welche Adapter zur RACE-Amplifikation ligiert wurden (MarathonTM-cDNA Amplification Kit, Clontech Laboratories GmbH, Tullastraße 4, 69126 Heidelberg, Deutschland). Zunächst wurde mittels spezifischer Primer (Nr. 121234 + Nr. 121236, Tabelle 6) ein PCR-Fragment mit einer Länge von etwa 800bp (Basenpaaren) amplifiziert (PCR-Programm: (Taq60-60)). Dieses wurde kloniert und sequenziert (Taq-Dye-Deoxy-Terminator Sequenzierungs-Kit, Applied Biosystems, Foster: City, CA USA) Brunnenweg 13, Weilderstadt). Die Sequenzierung des PCR-Produkts ergab eine Sequenz, die eine hohe Homologie zu der DNA-Sequenz von AHV-Sema aufweist, identisch zu der Sequenz der beiden ESTs.

35 Ein PCR-Fragment von 600bp wurde mit dem Primerpaar (Nr. 121237 + Nr. 121239, Tabelle 6) identifiziert. Es zeigte sich, das es sich um Klone mit DNA-Sequenzen des selben Gens handelte.

Beispiel 2:

40 Das 800bp PCR-Fragment aus Beispiel 1 wurde radioaktiv markiert (Random-Priming nach der Methode von [Feinberg (1983) Anal. Biochem. 132:6-13], mit ³²P-α-dCTP) und als Sonde für einen Multi-Tissue Northern-Blot (Human Multiple Tissue Northern Bot II, Clontech, Heidelberg, Germany), der mRNA-Proben aus den Geweben Milz, Thymus, Prostata, Testis, Ovarien, Dünndarm, Dickdarm und Leukozyten (PBL) enthält, verwendet. Dabei zeigte sich deutlich die Expression einer mRNA mit einer Länge von etwa 3.3kb in Milz und Keimdrüsen (Hoden, Eierstöcken), sowie schwächer in Thymus und Darm. Eine Hybridisierung eines Master-Blots (Dotblot mit RNA aus zahlreichen Geweben (Human RNA Master BlotTM, Clontech) bestätigte dieses Ergebnis und zeigte auch eine starke Expression in Plazenta-Gewebe.

45 Die Hybridisierung wurde für 16 Stunden unter stringenten Bedingungen (5xSSC, 50 mM Na-Phosphat pH 6.8, 50 % Formamid, 100 µg/ml Hefe-RNA) bei 42°C durchgeführt. Die Blots wurden stringent gewaschen (65°C, 0,2xSSC, 0.1 % SDS) und einem Fuji BAS2000 PhosphoimagerTM exponiert

Beispiel 3:

55 Eine cDNA-Bibliothek aus humaner Milz, kloniert in dem Bakteriophagen Lambda gt10 (Human Spleen 5' STRETCH PLUS cDNA, Clontech) wurde mit dieser Sonde durchsucht und ein Lambda-Klon identifiziert. Die in diesem Klon inserierte cDNA mit einer Länge von 1.6kb wurde mittels PCR (ExpandTM Long Template PCR System, Boehringer Mannheim GmbH, Sandhofer Straße 116, 68305 Mannheim) amplifiziert wobei die vektorspezifischen Primer Nr. 207608 + Nr. 207609 (Tabelle 6) verwendet wurden (flankierend der EcoRI-Klonierungsstelle) und das erhaltene PCR-

Fragment sequenziert. Dieser Klon enthielt das 5' Ende der cDNA und erweiterte die bekannte cDNA Sequenz auch nach 3'. Ausgehend von den neuen Teilsequenzen der cDNA wurden neue Primer für die RACE-PCR entwickelt (Nr. 232643, Nr. 232644, Nr. 233084, Tabelle 6). Zusammen mit einer verbesserten Thermocyclertechnik (PTC-200 von MJ-Research, Biozym Diagnostik GmbH, 31833 Hess. Oldendorf) mit deutlich besseren Leistungsdaten (Heiz- und Kühlrate) wurde ein 3'-RACE PCR-Produkt amplifiziert wobei die Primer Nr. 232644 bzw. Nr. 232643 und AP1 verwendet wurden und in den Vektor pCR2.1 (Invitrogen, De Schelp 12, 9351 NV Leek, Niederlande) kloniert. Das 3'-RACE PCR-Produkt wurde sequenziert und auf diese Weise das 3' Ende der cDNA identifiziert. Eine RACE-Amplifikation nach 5' (Primer Nr. 131990 bzw. Nr. 233084 und AP1) erweiterte das 5' Ende der cDNA um wenige Nukleotide und bestätigte den im identifizierten Lambda-Klon gefundenen Aminoterminus des H-SemaL.

Beispiel 4:

Ausgehend von einem kurzen murinen EST (Zuordnungs-Nr. AA260340) und einem daraus abgeleiteten Primer Nr. 260813 (Tabelle 6) und dem H-SemaL spezifischen Primern Nr. 121234 (Tabelle 6) wurde mittels PCR (Bedingungen: Taq52-60) ein DNA-Fragment mit einer Länge von ca. 840 bp muriner cDNA amplifiziert und in den Vektor pCR2.1 kloniert. Das dieses DNA-Fragment enthaltende Gen wurde M-SemaL genannt. Mit dem erhaltenen M-SemaL DNA-Fragment eine cDNA-Bank aus Mäuse-Milz (Mouse Spleen 5' STRETCH cDNA, Clontech) untersucht, wobei bereits mehrere Klone identifiziert werden konnten.

PCR (Taq60-30) mit den Primern Nr. 260812, Nr. 260813 aus muriner, endothelialer cDNA lieferte ein PCR-Fragment mit einer Länge von 244 Basenpaaren. Die PCR-Ergebnisse zeigten, daß eine deutliche basale Expression in murinen Endothelzellen vorhanden ist, welche nach Stimulation mit dem Zytokin Interferon- γ und Lipopolysacchariden zurückgeht.

Beispiel 5:

Untersuchungen zur Chromosomalen Lokalisation wurden mittels Fluoreszenz in-situ Hybridisierung (FISH) durchgeführt. Dazu wurden Metaphase Chromosomen von Mensch und Maus, ausgehend von einer humanen Blutprobe bzw. der Mauszelllinie BINE 4.8 (Keyna et al. (1995) J. Immunol. 155, 5536-5542), hergestellt (Kraus et al. (1994) Genomics 23, 272-274). Die Objektträger wurden mit RNase und Pepsin behandelt (Liehr et al. (1995) Appl. Cytogenetics 21, 185-188). Für die Hybridisierung wurden 120 mg humane Nick-translatierte Semaphorin-Probe bzw. 200 mg einer entsprechenden Mausprobe verwendet. Die Hybridisierung wurde jeweils in Gegenwart von 4.0 μ g COT1-DNA und 20 μ g STD bei 37°C (3 Tage) in einer befeuchteten Kammer durchgeführt.

Die Objektträger wurden mit 50% Formamid/2x SSC (3 mal je 5 min bei 45°C) und dann mit 2 x SSC (3 mal je 5 min bei 37°C) gewaschen und die biotinylierte Probe mit dem FITC-Avidin-System (Liehr et al. (1995)), detektiert. Die Objektträger wurden mit Hilfe eines Fluoreszenz-Mikroskops ausgewertet. Es wurden 25 Metaphasen/Probe ausgewertet, wobei jedes Experiment doppelt durchgeführt wurde. Es zeigt sich, daß H-SemaL auf Chromosom 15q23 lokalisiert ist Chromosomal benachbart liegen der Locus für das Bardet-Biedl-Syndrom und Tay-Sachs Erkrankung (Hexosaminidase A).

Beispiel 6:

Die genomische Intron-Exon Struktur des H-SemaL Gens ist zum größten Teil aufgeklärt.

Genomische DNA Fragmente wurden ausgehend von 250 mg humaner genomischer DNA, die aus PHA stimulierten peripheren Lymphozyten (Blut) isoliert worden waren, amplifiziert. Kürzere Fragmente wurden mit Ampli Taq^R (Perkin Elmer), längere Fragmente mit dem Expanded Long Templat PCR System^R (Boehringer Mannheim) amplifiziert.

Durch PCR-Amplifikation konnte bisher fast der vollständige genomische Locus des H-SemaL kloniert und charakterisiert werden. Insgesamt konnten bereits mehr als 8888 bp der genomischen Sequenz bestimmt werden und so die Intron-Exon-Struktur des Gens weitgehend aufgeklärt werden.

Beispiel 7:

Expressionsklonierungen:

Da kein kompletter Klon des Semaphoringens aus der Lambda-gt10 cDNA-Bank isoliert werden konnte, und auch mittels PCR ein vollständiger Klon nicht zu erhalten war, wurde der kodierende Bereich der cDNA in 2 überlappenden Subfragmenten mittels PCR (XL62-6) mit Hilfe der Primer Nr. 240655 und Nr. 121339 für das N-terminale DNA-Fragment, sowie den Primern Nr. 240656 (enthält HindIII und PmeI Schnittstellen) und Nr. 121234 für das C-terminale DNA-Fragment amplifiziert. Die erhaltenen DNA-Fragmente (Subfragmente) wurden in den Vektor pCR21 kloniert. Die beiden

Subfragmente wurden komplett sequenziert und schließlich die vollständige H-SemaL cDNA durch Insertion eines 0.6kb C-terminalen SstI-HindIII Restriktions-Fragments in das mit den Restriktionsenzymen SstI und HindIII geschnittene, das N-terminale DNA-Fragment enthaltende Plasmid, hergestellt. Aus diesem Plasmid pCR2.1-H-SemaL (gemäß Sequenz in Tabelle 7, SEQ ID NO. 34) wurde das komplette Gen mittels der EcoRI-Schnittstelle (in pCR2.1) und HindIII-Schnittstelle (in Primer Nr. 240656, Tabelle 6) herausgeschnitten und in einen entsprechend geschnittenen, konstitutiven Expressionsvektor pCDNA3.1(-)MycHisA (Invitrogen) ligiert. Aus dem resultierenden rekombinanten Plasmid pCDNA3.1(-)H-SemaL-MycHisA (gemäß Sequenz in Tabelle 8) wurde das EcoRI-ApaI Fragment (ohne Myc-His-Tag) herausgeschnitten und in den induzierbaren Vektor pIND ligiert (Ecdysone-Inducible Mammalian Expression System, Invitrogen), der zuvor ebenfalls mit EcoRI-ApaI geschnitten worden war. Das rekombinante Plasmid wurde mit pIND-H-SemaL-EA (Sequenz gemäß Tabelle 11) bezeichnet. Ein EcoRI-PmeI-Fragment (mit Myc-His-Tag) aus pCDNA3.1(-)H-SemaL-Myc-HisA (Sequenz gemäß Tabelle 9) wurde in einen mit EcoRI-EcoRV geschnittenen Vektor pIND eingesetzt. Das rekombinante Plasmid wurde mit pIND-H-SemaL-EE (Sequenz gemäß Tabelle 10) bezeichnet. Ein Fusionsgen von H-SemaL mit Enhanced Green Fluorescent Protein (EGFP) wurde hergestellt durch Ligation des mit PCR amplifizierten EGFP-Leserahmens (aus dem Vektor pEGFP-C1 (Clontech), mit Hilfe der Primer Nr. 243068 + Nr. 243069, Taq52-60) in die PmeI-Schnittstelle des Plasmids pCDNA3.1(-)H-SemaL-MycHisA wodurch das Plasmid pCDNA3.1(-)H-SemaL-EGFP-MycHisA (Sequenz gemäß Tabelle 9) erhalten wurde.

In den Tabellen 7 bis 13 bedeuten kleine Buchstaben die Sequenz von H-SemaL, Teilen oder Derivaten derselben und große Buchstaben die Sequenz des Plasmids.

Beispiel 8:

Zur Herstellung von H-SemaL spezifischen Antikörpern wurden cDNA-Fragmente von H-SemaL in prokaryotische Expressionsvektoren integriert, in E. coli exprimiert und die Semaphorin-Derivate aufgereinigt. Die Semaphorin-Derivate wurden als Fusionsproteine mit einem His-Tag exprimiert. Dementsprechend wurden Vektoren verwendet, die die Sequenz für ein His-Tag enthalten und eine Integration des Semaphorin cDNA-Fragments im Leserahmen ermöglichen. Ein N-terminales 6xHistidin-Tag ermöglicht z.B. eine Aufreinigung mittels Nickel-Chelat-Affinitätschromatographie (Qiagen GmbH, Max-Volmer Straße 4, 40724 Hilden):

1. Der für die Aminosäuren 179-378 kodierende Teil der H-SemaL cDNA wurde mittels PCR mit den Primern Nr. 150788 und Nr. 150789 amplifiziert und dieses DNA-Fragment in den Vektor pQE30 (Qiagen), der zuvor mit den Restriktionsenzymen BamHI und HindIII geschnitten worden war, ligiert (Konstrukt pQE39-H-SemaL-BH (Sequenz gemäß Tabelle 12)).

2. Der für die C-terminalen Aminosäuren 480-666 kodierende Abschnitt der H-SemaL cDNA wurde mit den Restriktionsenzymen SstI und HindIII aus dem Plasmid pCR 2.1 geschnitten und in den Vektor pQE31 (Qiagen), der zuvor mit SstI und HindIII geschnitten worden war ligiert (Konstrukt pQE31-H-SemaL-SH (Sequenz gemäß Tabelle 13)).

Die korrekte Integration der Sequenzen im richtigen Leserahmen wurde durch DNA-Sequenzierung überprüft. Die Fusionsproteine, bestehend aus einem N-terminalen 6xHistidin-Tag und einem Teil des Semaphorins H-SemaL wurden mittels Ni^{2+} -Affinitätschromatographie aufgereinigt. Die aufgereinigten Fusionsproteine wurden zur Immunisierung von verschiedenen Tieren (Hase, Huhn, Maus) benutzt.

Beispiel 9:

FACS-Analyse verschiedener Zelltypen (Figuren 4 und 5)

Die Zellen (ca. $0.2-0.5 \times 10^6$) wurden mit FACS-Puffer gewaschen (Phosphate-buffered Saline (PBS) mit 5% fötalem Kälberserum (FCS) und 0.1% Na-Azid) und dann jeweils (auf Eis) für 1 Stunde mit den Antiseren inkubiert.

Als primäre Antikörper dienten für die Kontrolle (Overlay Hühner-Präimmunserum (1 : 50) und für den spezifischen Nachweis (Spezifische Färbung) ein H-SemaL spezifisches Hühner-Antiserum (1 : 50).

Das spezifische Antiserum mit Antikörpern gegen Aminosäuren (Aa) 179-378 (mit N-terminalem His-Tag) von H-SemaL wurde durch Immunisierung von Hühnern mit dem durch Ni-Chelat-Affinitätschromatographie gereinigten Protein erzeugt (wie in Beispiel 8 beschrieben).

Als zweiter Antikörper wurde ein FITC-markierter anti-Huhn F(ab') Antikörper aus Kaninchen verwendet (Dianova Jackson Laboratories, Best.-Nr. 303-095-006, Hamburg, Deutschland) (1mg/ml).

Für die CD100-Färbung: wurde ein Rabbit-anti-Maus IgG, FITC-markiert verwendet. Der zweite Antikörper wurde jeweils in 1 : 50 Verdünnung in FACS-Puffer eingesetzt.

Dann wurden die Zellen gewaschen, in PBS resuspendiert und im FACS analysiert. Die FACS-Analyse wurde mit

einem FACS-Trak Gerät (Becton—Dickinson) durchgeführt Prinzip: Eine Einzelzellsuspension wird in einen Messkanal vorbeigeleitet, dort werden die Zellen mit Laserlicht von 488nm bestrahlt und so Fluoreszenzfarbstoffe (FITC) angeregt. Gemessen werden Streulicht nach vorne (forward scatter, FSC: korreliert mit der Zellgröße), zur Seite (sideward scatter, SSC: korreliert mit dem Granulargehalt: bei unterschiedlichen Zelltypen unterschiedlich) und Fluoreszenz im Kanal 1 (FL 1) (für Wellenlängen im FITC Emissionsbereich, max. bei 530nm). Auf diese Weise wurden je 10000 Ereignisse (Zellen) gemessen.

Der Dotplot (Figuren 4a - k) (jeweils linke Abbildung): FSC gegen SSC (Größe gegen Granulargehalt/Streuung), darin eingegrenzt ist die im rechten Fenster (jeweils zugehörige rechte Abbildung) analysierte (einheitliche) Zellpopulation von ähnlicher Größe und Granulargehalt. Das rechte Fenster zeigt die Intensität von FL1 (X-Achse) gegen die Zahl der Ereignisse (Y-Achse), also eine Häufigkeitsverteilung.

Hierbei ist jeweils das Ergebnis mit dem Kontrollserum (nichtgefüllte Kurve) dem Ergebnis der spezifischen Färbung (ausgefüllte Kurve) überlagert. Eine Verschiebung der Kurve für die spezifische Färbung gegenüber der Kontrolle nach rechts entspricht einer Expression von H-SemaL in den entsprechenden Zellen. Je weiter die Verschiebung, desto stärker ist die Expression.

15 Für FACS Analyse verwendete Zelllinien:

a) Zelllinie U937

American Type Culture Collection ATCC; ATCC Nummer: CRL-1593 Name: U-937

Gewebe: lymphoma; histiocytic; Monozytenähnlich

20 Species: human;

Hinterleger: H. Koren

b) Zelllinie THP-1

ATCC Nummer: TIB-202

Gewebe: monocyte; acute monocytic leukemia

25 Species: human

Hinterleger: S. Tsuchiya

c) Zelllinie K-562

ATCC Nummer: CCL-243

Gewebe: chronic myelogenous leukemia

30 Species: human;

Hinterleger H.T. Holden

d) Zelllinie L-428

DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, DSMZ Nr: ACC 197

Zelltyp: human Hodgkin's lymphoma

35 e) Zelllinie Jurkat

DSMZ-Deutsche Sammlung von Mikroorganismen und zellkulturen GmH, DSMZ Nr: ACC 282

Zelltyp: human T cell leukemia

f) Zelllinie Daudi

ATCC Nummer: CCL-213

40 Gewebe: Burkitt's lymphoma; B lymphoblast; B Zellen

Species : human

Hinterleger: G. Klein

g) Zelllinie LCL

EBV-transformierte lymphoblastoide B-Zelllinie.

45 h) Zelllinie Jiyoye (P-2003)

ATCC Nummer: CCL-87

Gewebe: Burkitt's lymphoma; B Zellen; B lymphocyte

Species: human

Hinterleger: W. Henle

50 i) CBL-Mix57

humane T-Zelllinie (isoliert aus Blut) transformiert mit rekombinantem H. Saimiri (Wild-typ ohne Deletion)

j) CBL-Mix59

humane T-Zelllinie (isoliert aus Blut) transformiert mit H. Samiri (Deletion von ORF71).

55 Beispiel 10: Proteingel und Western-Blot

Sekretierbare humane SemaL (Aminosäuren 42-549 in Tabelle 4 (ohne Signalpeptid und ohne Transmembrandomäne) wurde in das Plasmid pMelBac-A (Invitrogen, De Sehelp, Leck, Niederlande, Cv 1950-20) kloniert und auf diese

Weise das Plasmid pMelBacA-H-SemaL (Länge 6622bp) erzeugt (Tabelle 15, Figur 8). Das H-SemaL Derivat wurde im Baculovirus-System (Bac-N-Blue, Invitrogen) exprimiert. Die Expression wurde in den Insekten-Eizellen abgeleiteten Zelllinien Sf9 (von *Spodoptera flugiperda*) und High Five™ (von *Trichoplusia ni*, U.S. Pat. No. 5,300,435, gekauft von Invitrogen) durch Infektion mit den rekombinanten, plaquegereinigten Baculoviren ausgeführt.

Die Expression wurde nach den Angaben des Herstellers durchgeführt. Anschließend wurden die Proteine in einem Gel aufgetrennt und das H-SemaL Derivat im Western-Blot nachgewiesen. Die Detektion wurde mit dem H-SemaL spezifischen Hühnerantiserum (vergl. Beispiel 8 und Figur 7) (Verdünnung: 1:100) durchgeführt. Der spezifische Hühnerantikörper wurde mit anti-IgY-HRP Konjugat (Verdünnung: 1:3000, vom Kaninchen; Dianova Jackson Laboratories) nach Angaben des Herstellers nachgewiesen.

Beispiel 11: Herstellung von pMelBacA-HSEMA

Der rekombinante Vektor (pMelBacA-HSEMA, 6622bp) wurde hergestellt, indem ein entsprechendes DNA-Fragment, das für die Aminosäuren 42-649 von H-SemaL kodiert, in den Vektor pMelBacA (4,8 kb, Invitrogen) kloniert wurde (vergl. Annotation zu pMelBacA-H-SemaL). Die Klonierung erfolgte über BamHI und EcoRI in frame hinter die in dem Vektor vorliegende Signalsequenz („honeybee melittin signal sequence“). Ein entsprechendes H-SemaL DNA-Fragment wurde mit dem Primerpaar H-Sema 1 baculo 5' und H-Sema 1 baculo 3' amplifiziert.

Primer zur Amplifikation (TaKaRa Ex Taq-Polymerase) und Klonierung:

„H-Sema 1 baculo 5'“, zur Amplifikation ohne Signalsequenz und zur Einführung einer BamHI-Schnittstelle

5'-CCGGATCCGCCAGGGCCACCTAAGGAGCGG-3' (SEQ ID NO: 43)

„H-Sema 1 baculo 3'“, zur Amplifikation ohne Transmembrandomäne und zur Einführung einer EcoRI-Schnittstelle

5'-CTGAATTCAGGAGCCAGGGCACAGGCATG-3' (SEQ ID NO: 44).

Abbildungen:

Figur 1:

Gewebespezifische Expression von H-SemaL

A) Mehrfach-Gewebe Northern Blot (Clontech, Heidelberg, Deutschland). Von links nach rechts sind aufgetragen: Je 2 µg Poly-A-RNA pro Spur aus Milz, Thymus, Prostata, Hoden, Eierstöcken, Dünndarm, Dickdarm Mukosa, peripheren (Blut-) Leukozyten. Größenstandards sind markiert.

Die Blots wurden unter stringenten Bedingungen mit einer 800 Basenpaaren langen H-SemaL Probe hybridisiert.

Figur 2:

Schematische Darstellung der Klonierung der H-SemaL cDNA und der genomischen Organisation der H-SemaL kodierenden Sequenzen (H-SemaL Gen)

Oben: Lokalisation der EST-Sequenzen (Zugangsnummern; Lage der EST-Sequenzen ist relativ zur AHV-Sema Sequenz dargestellt).

Darunter: Amplifizierte PCR- und RACE-Produkte sowie die Position der cDNA Klone im Bezug auf die Lokalisation in der vollständigen H-SemaL cDNA und dem offenen Leserahmen (ORF) für das kodierte Protein.

Unten: Relative Position der Exons im H-SemaL Gen im Bezug auf die genomische Sequenz.

Die Position der verwendeten Oligonukleotid Primer ist durch Pfeile angezeigt

Figur 3:

Phylogenetischer Baum: Erhalten durch mehrfaches Alignment der aufgeführten Semaphorin Sequenzen. Aufgrund der Gruppierung der Semaphorine in dem phylogenetischen Baum kann auf deren phylogenetische Verwandtschaft geschlossen werden.

Figur 4:

FACS Analyse der H-SemaL Expression in verschiedenen Zelllinien bzw. verschiedenen Zelltypen (vergl. Beispiel 8).

Figur 5:

Vergleichende Analyse der CD 100 und H-SemaL Expression (vergl. Beispiel 9).

5 Figur 6: Expression von sekretierbarem humanen SEMAL (H-SemaL) in HiFive und Sf3-Zellen (vergl. Beispiel 10).

(Aa 42-649 in pMelBac-A (Invitrogen) im Baculovirus-System (Bac-N-Blue, Invitrogen) Detektion mit spezifischem Hühner-Antiserum (1:100) und anti-IgY-HRP Konjugat (1:3000, vom Kaninchen, Jackson Lab.)

1,4,6 uninfizierte HiFive Zellen (serumfrei)

10 2,3,5,7,8 mit rekombinantem Baculovirus infizierte HiFive Zellen (serumfrei)

M Rainbow molecular weight marker (Amersham RPN756)

9,10 infizierte Sf9 Zellen (serumhaltiges Medium).

Figur 7: Spezifität des Antiserums

15

Spuren 1-3: Huhn 1; Spuren 4-6: Huhn 2

Spuren 1 und 4: Präimmunserum

Spuren 2 und 5: 60. Immunisierungstag

Spuren 4 und 6: 105. Immunisierungstag

20 Immunisiert wurde mit den Aminosäuren 179-378 von H-SemaL (mit aminoterminalen His-Tag) (vergl. Beispiel 8, Punkt 1.)

Figur 8: Abbildung der Plasmidkarte von pMelBacA-HSEMAL.

25 Das rekombinante Plasmid wurde wie in Beispiel 11 beschrieben, hergestellt.

30

35

40

45

50

55

Tabelle 1: Verschiedene Subtypen von Semaphorinen aus verschiedenen Spezies

| Name | Synonym | Spezies | | Referenz |
|------------|-------------|---------|---|--|
| H-Sema III | (H-SemaD) | Mensch | Sez. | (Kolodkin et al. 1993) |
| CD-100 | | Mensch | TM, IC; CD45 assoziiert, in T-Zellen exprimiert | (Hall et al. 1996) |
| H-Sema V | (H-SemaA) | Mensch | Sez.; Locus 3p21.3 | (Sekido et al. 1996; Roche et al. 1996) |
| H-Sema IV | (H-Sema 3F) | Mensch | Sez.; Locus 3p21.3 | (Xiang et al. 1996; Sekido et al. 1996) |
| H-SemaE | | Mensch | Sez.; am 3'Ende von M-SemaE divergierend (Leserahmen im Alignment verbessert) | AB000220 (Yamada 1997 unpublished) |
| H-SemaK | KIAA0331 | Mensch | Sez.; | (Nagase et al. 1997) |
| H-SemaL | SEMAL | Mensch | TM, kein IC | Diese Anmeldung |
| M-SemaA | | Maus | Sez. | (Püschel et al. 1995) |
| M-SemaB | | Maus | TM, IC | (Püschel et al. 1995) |
| M-SemaC | | Maus | TM, IC | (Püschel et al. 1995) |
| M-SemaD | M-Sema III | Maus | Sez. | (Messersmith et al. 1995; Püschel et al. 1995) |
| M-SemaE | | Maus | Sez.; 5 partielle Sequenz | (Püschel et al. 1995) |
| M-SemaF1 | M-SemaF | Maus | TM, IC | (Inagaki et al. 1995) |
| M-SemaG2 | M-SemaG | Maus | TM, IC; exprimiert in Lymphoiden Zellen, Maus-Homolog zu CD100 | (Funayama et al. 1996) |

| Name | Synonym | Spezies | | Referenz |
|-------------|-------------|--------------------|------------------------------|---------------------------------|
| M-Sema-F2 | M-Sema-F | Maus | TM, IC; Thrombospondin-Motiv | (Adams et al. 1996) |
| M-SemaG1 | M-SemaG | Maus | TM, IC; Thrombospondin-Motiv | (Adams et al. 1996) |
| M-SemaH | | Maus | Sez. | (Christensen 1996 unpub) Z80941 |
| M-SemaVia | | Maus | TM, IC | (Zhou et al. 1997) |
| M-SemaL | SemaL | Maus | Partielle Sequenz | Diese Anmeldung |
| Collapsin-1 | | Huhn | Sez. | (Luo et al. 1993) |
| Collapsin-2 | | Huhn | Sez. | (Luo et al. 1995) |
| Collapsin-3 | | Huhn | Sez. | (Luo et al. 1995) |
| Collapsin-4 | | Huhn | Partielle Sequenz | (Luo et al. 1995) |
| Collapsin-5 | | Huhn | Sez. | (Luo et al. 1995) |
| R-Sema III | | Ratte | Sez. | (Giger et al. 1996) |
| T-Sema I | | Tribolium confusum | TM, IC | (Kolodkin et al. 1993) |
| Ce-Sema I | | C.elegans | TM, IC | U15667 (Roy 1994 unpublished) |
| G-Sema I | Fascidin-IV | Grashüpfer | TM, IC | (Kolodkin et al. 1992) |
| D-Sema I | | Drosophila | TM, IC | (Kolodkin et al. 1993) |
| D-Sema II | | Drosophila | Sez. | (Kolodkin et al. 1993) |
| AHV-Sema | | AHV-1 | Sez. | (Ensner and Fleckenstein, 1995) |

5

10

15

20

25

30

35

40

45

50

55

| Name | Synonym | Spezies | | Referenz |
|-----------------|---------|----------|-------|------------------------|
| ORF-A39 | | Vaccinia | Sez. | (Kolodkin et al. 1993) |
| ORF-A39-homolog | | Variola | Sez.; | (Kolodkin et al. 1993) |

TM: Transmembrandomäne

Sez: sezerniert

IC: vermutlich intrazelluläres zytoplasmatisches Sequenz-Motiv

Tabelle 2: cDNA-Sequenz von H-SemaL (2636 Nukleotide) (SEQ ID NO.: 1)

```

5      1      cggggccacg ggatgacgcc tcctccgccc ggacgtgccg cccccagcgc
51     accgcgcgcc cgcgtccctg gccgcgcggc tcggttgggg cttccgctgc
101    ggctgcggct gctgctgctg ctctgggcgg ccgcgcctc cggccagggc
151    cacctaagga gcggaacccg catcttcgcc gtctggaaag gccatgtagg
10     201    gcaggaccgg gtggactttg gccagactga gccgcacacg gtgcttttcc
251    acgagccagg cagctcctct gtgtgggtgg gaggacgtgg caaggtctac
301    ctctttgact tccccgaggg caagaacgca tctgtgcgca cgggtgaatat
15     351    cggctccaca aaggggtcct gtctggataa gcgggactgc gagaactaca
401    tcactctcct ggagaggcgg agtgaggggc tgctggcctg tggcaccaac
451    gcccggcacc ccagctgctg gaacctggtg aatggcactg tggtgccact
20     501    tggcgagatg agaggctacg ccccttcag cccggacgag aactccctgg
551    ttctgtttga aggggacgag gtgtattcca ccatccggaa gcaggaatac
601    aatgggaaga tccttcggtt ccgccgcac cggggcgaga gtgagctgta
651    caccagtgat actgtcatgc agaaccaca gttcatcaaa gccaccatcg
25     701    tgcaccaaga ccaggcttac gatgacaaga tctactactt cttccgagag
751    gacaatcctg acaagaatcc tgaggctcct ctcaatgtgt cccgtgtggc
801    ccagttgtgc aggggggacc agggtgggga aagttcactg tcagtctcca
851    agtggaacac ttttctgaaa gccatgctgg tatgcagtga tgctgccacc
30     901    aacaagaact tcaacaggct gcaagacgtc ttctgtctcc ctgaccccag
951    cggccagtgg agggacacca gggctctatgg tgttttctcc aacctctgga
1001   actactcagc cgtctgtgtg tattccctcg gtgacattga caaggtcttc
35     1051   cgtacctcct cactcaaggg ctaccactca agccttccca acccgcgggc
1101   tggcaagtgc ctcccagacc agcagccgat acccacagag accttccagg
1151   tggctgaccg tcaccagag gtggcgaga gggtgaggcc catggggcct
1201   ctgaagacgc cattgttcca ctctaaatac cactaccaga aagtggccgt
40     1251   tcaccgcatg caagccagcc acggggagac ctttcatgtg ctttacctaa
1301   ctacagacag gggcactatc cacaaggtgg tggaaccggg ggagcaggag
1351   cacagcttcg cttcaacat catggagatc cagcccttcc gccgcgcggc
45     1401   tgccatccag accatgtcgc tggatgctga gcggaggaag ctgtatgtga
1451   gctcccagtg ggaggtgagc caggtgcccc tggacctgtg tgaggtctat
1501   ggcgggggct gccacggtt cctcatgtcc cgagaccctt actgcggctg
1551   ggaccagggc cgctgcatct ccactacag ctccgaacgg tcagtgtcgc
50     1601   aatccattaa tccagccgag ccacacaagg agtgtcccaa ccccaaacca
1651   gacaaggccc cactgcagaa ggtttccctg gcccctaaact ctgcgtacta
1701   cctgagctgc cccatggaat cccgccacgc cacctactca tggcgccaca

```

5

10

15

20

25

1751 aggagaacgt ggagcagagc tgcgaacctg gtcaccagag ccccaactgc
 1801 atcctgttca tcgagaacct cacggcgcag cagtacggcc actacttctg
 1851 cgaggcccag gagggtcctt acttccgcga ggctcagcac tggcagctgc
 1901 tgcccagga cggcacatg gccgagcacc tgctgggtca tgccctgtgcc
 1951 ctggctgcct cctcttggtt ggggtgctg cccacactca ctcttggtt
 2001 gctggtccac tagggcctcc cgaggctggg catgcctcag gcttctgcag
 2051 cccagggcac tagaacgtct cactcaga gccggctggc ccgggagctc
 2101 cttgcctgcc acttcttcca gggacagaa taaccagtg gaggatgcca
 2151 ggcctggaga cgtccagccg caggcgctg ctgggcccc ggtggcgcac
 2201 ggatggtgag gggctgagaa tgagggcacc gactgtgaag ctggggcacc
 2251 gatgaccaa gactttatct tctggaaaat atttttcaga ctcccaaac
 2301 ttgactaaat gcagcgatgc tccagccca agagcccatg ggtcggggag
 2351 tgggtttgga taggagagct gggactccat ctgacctg gggctgaggc
 2401 ctgagtcctt ctggactctt ggtaccaca ttgcctcctt cccctccctc
 2451 tctcatggct ggggtgctgg tgttctgaa gaccagggc taccctctgt
 2501 ccagccctgt cctctgcagc tccctctctg gtcctgggtc ccacaggaca
 2551 gccgccttgc atgtttattg aaggatgttt gctttccgga cggaaggacg
 2601 gaaaaagctc tgaaaaaaaa aaaaaaaaaa aaaaaa

Tabelle 3: Nukleotidsequenz der cDNA von M-SemaL (partiell, 1195 Nukleotide)

(SEQ ID NO.: 2)

30

35

40

45

50

55

1 cggggctgcg ggatgacgcc tctctctccc ggacgtgccg cccccagcgc
 51 accgcgcgcc cgcgtcctca gcctgccggc tcggttcggg ctcccgctgc
 101 ggctgcccgt tctgctggtg ttctgggtgg ccgccgcctc cgcccaaggc
 151 cactcgagga gcggaccccg catctccgcc gtctggaaaag ggcaggacca
 201 tgtggacttt agccagcctg agccacacac cgtgcttttc catgagccgg
 251 gcagcttctc tgtctgggtg ggtggacgtg gcaaggtcta ccacttcaac
 301 ttccccgagg gcaagaatgc ctctgtgcgc acggtgaaca tcggctccac
 351 aaaggggtcc tgtcaggaca aacaggactg tgggaattac atcactcttc
 401 tagaaaggcg gggtaatggg ctgctggtct gtggcaccaa tgcccgaag
 451 cccagctgct ggaacttggg gaatgacagt gtggtgatgt cacttgggtg
 501 gatgaaaggc tatgccccct tcagcccga tgagaactcc ctggttctgt
 551 ttgaaggaga tgaagtgtac tctaccatcc ggaagcagga atacaacggg
 601 aagatccctc ggtttcgacg cattcggggc gagagtgaac tgtacacaag
 651 tgatacagtc atgcagaacc cacagttcat caaggccacc attgtgcacc
 701 aagaccaagc ctatgatgat aagatctact acttcttccg agaagacaac
 751 cctgacaaga accccgaggc tcctctcaat gtgtcccgag tagcccagtt

801 gtgcaggggg gaccaggggtg gtgagagttc gttgtctgtc tccaagtgga
 851 acaccttcct gaaagccatg ttgggtctgca gcgatgcagc caccaacagg
 5 901 aacttcaatc ggctgcaaga tgtcttcctg ctccctgacc ccagtggcca
 951 gtggagagat accaggggtct atggcgtttt ctccaacccc tggaactact
 1001 cagctgtctg cgtgtattcg cttgggtgaca ttgacagagt cttccgtacc
 1051 tcatcgctca aaggctacca catgggcctt tccaaccctc gacctggcat
 1101 gtgcctccca aaaaagcagc ccatacccac agaaaccttc caggtagctg
 1151 atagtcaccc agaggtggct cagaggggtg aacctatggg gcccc

15 Tabelle 4: Aminosäuresequenz von H-SemaL (666 Aminosäuren)
 (SEQ ID NO.: 3)

20 1 MTPPPPGRAA PSAPRARVPG PPARLGLPLR LRLLLLWAA AASAQGHLS
 51 GPRIFAVWKG HVGQDRVDFG QTEPHTVLFH EPGSSSVWVG GRGKVYLFDF
 101 PEGKNASVRT VNIGSTKGSC LDKRDCENYI TLLERRSEGL LACGTNARHP
 151 SCWNLVNGTV VPLGEMRGYA PFSPDENSIV LFEGDEVYST IRKQEYNGKI
 25 201 PRFRRIRGES ELYTSDTVMQ NPQFIKATIV HQDQAYDDKI YYFFREDNPD
 251 KNPEAPLNVS RVAQLCRGDQ GGESSLSVSK WNTFLKAMLV CSDAATNKNF
 301 NRLQDVFLLP DSPGQWRDTR VYGVFSNPWN YSAVCVYSLG DIDKVFTSS
 351 LKGYHSSLPN PRPGKCLPDQ QPIPTETFQV ADRHPEVAQR VEPMGPLKTP
 30 401 LFHSKYHYQK VAVHRMQASH GETFHVLYLT TDRGTIHKVV EPGEQHSFA
 451 FNIMEIQPFR RAAAIQTMSL DAERRKLYVS SQWEVSQVPL DLCEVYGGGC
 501 HGCLMSRDY CGWDQGRCS IYSSERSVLQ SINPAEPHKE CPNPKPKAP
 551 LQKVS LAPNS RYLLSCPMES RHATYSWRHK ENVEQSCEPG HQSPNCILFI
 35 601 ENLTAQQYGH YFCEAQEGSY FREAQHWQLL PEDGIMAEHL LGHACALAAS
 651 LWLGVLPTLT LGLLVH

40 Tabelle 5: (Partielle) Aminosäuresequenz von M-SemaL (394 Aminosäuren, entspricht Position 1-396 von
 H-SemaL) (SEQ ID NO.: 4)

45 1 MTPPPPGRAA PSAPRARVLS LPARFGLPLR LRLLLVFWA AASAQGHRS
 51 GPRISAVWKG QDHVDFSQPE PHTVLFHEPG SFSVWVGGRG KVYHFNFPFG
 101 KNASVRTVNI GSTKGSCQDK QDCGNYITLL ERRGNGLLVC GTNARKPSCW
 151 NLVNDVVMV LGEMKGYAPF SPDENSIVLF EGDEVYSTIR KQEYNGKIPIR
 50 201 FRRIRGESEL YTSDTVMQNP QFIKATIVHQ DQAYDDKIYY FFREDNPKN
 251 PEAPLNVS RV AQLCRGDQGG ESSLSVSKWN TFLKAMLVCS DAATNRNENR

301 LQDVFLLPDP SGQWRDTRVY GVESNPWNYS AVCVYSLGDI DRVFTSSLK
 351 GYHMGLSNPR PGMCLPKKQP IPTETFQVAD SHPEVAQRVE PMGP

5

Tabelle 6: Synthetische Oligonukleotide, (Eurogentec, Seraing, Belgien)

10

Nummer des Nukleotidsequenz des Primers (der synthetischen
 Primers/Bezeichnung Oligonukleotide)

| | | |
|--------------|-------------------------|------------------|
| 91506/AP2 | actcactatagggctcgagcggc | (SEQ ID NO.: 5) |
| 121234 | agccgcacacggtgcttttc | (SEQ ID NO.: 6) |
| 121235/Est 2 | gcacagatgcgttcttgccc | (SEQ ID NO.: 7) |
| 121236/Est 3 | accatagaccctggtgtccc | (SEQ ID NO.: 8) |
| 121237/Est 4 | gcagtgatgctgccaccaac | (SEQ ID NO.: 9) |
| 121238 | ccagaccatgtcgtggatg | (SEQ ID NO.: 10) |

25

30

35

40

45

50

55

| | | | |
|----|--|--|------------------|
| | 121239/Est 6 | acatgaggcaaccgtggcag | (SEQ ID NO.: 11) |
| | 131989/AP1 | ccatcctaatacgaactcactatagggc | (SEQ ID NO.: 12) |
| 5 | 131990/Est 7 | aggtagaccttgccacgtcc | (SEQ ID NO.: 13) |
| | 131991 | gaacttcaacaggctgcaagacg | (SEQ ID NO.: 14) |
| | 131992 | atgctgagcggaggaagctg | (SEQ ID NO.: 15) |
| | 131993 | ccgccatacacctcacacag | (SEQ ID NO.: 16) |
| 10 | 150788 | ctggaagctttctgtgggtatcggctgc | (SEQ ID NO.: 17) |
| | 150789 | tttggatccctggttctgtttgaag | (SEQ ID NO.: 18) |
| | 167579/cDNA | ttctagaattcagcgccgctttttttttttttttttttttttttttttttttttttvn | |
| 15 | Synthese Primer | (SEQ ID NO.: 19) | |
| | 168421 | ggggaaagtctcactgtcagtcctcaag | (SEQ ID NO.: 20) |
| | 168422 | gggaatacacacagacggctgagtag | (SEQ ID NO.: 21) |
| | 207608/ | agcaagttcagcctggttaagt | (SEQ ID NO.: 22) |
| 20 | Amplifikation von λ gt10 Insert | | |
| | 207609/ | ttatgagtattttcttccagg | (SEQ ID NO.: 23) |
| | Amplifikation von λ gt 10 Insert | | |
| 25 | 232643/Est 13 | ccattaatccagccgagccacacaag | (SEQ ID NO.: 24) |
| | 232644/Est 14 | catctacagctccgaacggtcagt | (SEQ ID NO.: 25) |
| | 233084 | cagcgggaagccccaaccgag | (SEQ ID NO.: 26) |
| | 240655/hs 5 | gggatgacgcctcctccgcccgg | (SEQ ID NO.: 27) |
| 30 | 240656/hs 3 | aagcttcacgtggaccagcaagccaagagt | (SEQ ID NO.: 28) |
| | 240657/hs 3c | aagctttttcgcctcctccgcccgg | (SEQ ID NO.: 29) |
| | 243068 | atggtgagcaagggcgaggagctg | (SEQ ID NO.: 30) |
| | 243069 | cttgtagcagctcgtccatgccgag | (SEQ ID NO.: 31) |
| 35 | 260812 | GGGTGGTGAGAGTTCGTTGTCTGTC | (SEQ ID NO.: 32) |
| | 260813 | GAGCGATGAGGTACGGAAGACTCTG | (SEQ ID NO.: 33) |

Tabelle 7: Nukleotidsequenz des rekombinanten Plasmids pCR2.1-H-SemaL (SEQ ID NO.: 34)

| | | | | | | |
|----|-----|------------|------------|------------|------------|-------------|
| 40 | 1 | AGCGCCCAAT | ACGCAAACCG | CCTCTCCCCG | CGCGTTGGCC | GATTCATTAA |
| | 51 | TGCAGCTGGC | ACGACAGGTT | TCCCGACTGG | AAAGCGGGCA | GTGAGCGCAA |
| 45 | 101 | CGCAATTAAT | GTGAGTTAGC | TCACTCATTA | GGCACCCCAG | GCTTTACTACT |
| | 151 | TTATGCTTCC | GGCTCGTATG | TTGTGTGGAA | TTGTGAGCGG | ATAACAATTT |
| | 201 | CACACAGGAA | ACAGCTATGA | CCATGATTAC | GCCaagcttc | acgtggacca |
| | 251 | gcaagccaag | agtgagtgtg | ggcagcaccc | ccagccagag | ggaggcagcc |
| 50 | 301 | agggcacagg | catgaccag | caggtgctcg | gcatgatgc | cgctcctcggg |
| | 351 | cagcagctgc | cagtgtgag | cctcgcggaa | gtaggagccc | tcctgggcct |
| 55 | 401 | cgcagaagta | gtggccgtac | tgctgcgccg | tgaggttctc | gatgaacagg |

451 atgcagttgg ggctctggtg accaggttcg cagctctgct ccacgttctc
 501 cttgtggcgc catgagtagg tggcgtggcg ggattccatg gggcagctca
 551 ggtagtagcg agagtttggg gccagggaaa ccttctgcag tggggccttg
 601 tctggtttgg ggttgggaca ctccctgtgt ggctcggctg gattaatgga
 651 ttgcagcact gaccgttcgg agctgtagat ggagatgcag cggccctggt
 701 cccagccgca gtaggggtct cgggacatga ggcaaccgtg gcagcccccg
 751 ccatagacct cacacaggtc caggggcacc tggctcacct cccactggga
 801 gctcacatac agcttcctcc gctcagcatc cagcgacatg gtctggatgg
 851 cagccgcgcg gcggaagggc tggatctcca tgatgttgaa ggcgaagctg
 901 tgctcctgct cccccggttc caccaccttg tggatagtgc ccctgtctgt
 951 agttaggtaa agcacatgaa aggtctcccc gtggctggct tgcatgcggt
 1001 gaacggccac tttctggtag tggtatntag agtggaaaca tggcgtcttc
 1051 agaggcccca tgggctccac cctctgcgcc acctctgggt gacggtcagc
 1101 cacctggaag gtctctgtgg gtatcggctg ctggtctggg aggcacttgc
 1151 caggccgcgg gttgggaagg cttgagtgg agcccttgag tgaggaggta
 1201 cggaagacct tgtcaatgtc accgagggaa tacacacaga cggctgagta
 1251 gttccagggg ttggagaaaa caccatagac cctggtgtcc ctccactggc
 1301 cgctggggtc agggagcagg aagacgtctt gcagcctgtt gaagtctctg
 1351 ttggtggcag catcactgca taccagcatg gctttcagaa aagtgttcca
 1401 cttggagact gacagtgaac tttccccacc ctggtcccc ctgcacaact
 1451 gggccacacg ggacacattg agaggagcct caggattctt gtcaggattg
 1501 tcctctcgga agaagtagta gatcttgta tcgtaagcct ggtcttgggtg
 1551 cacgatggtg gctttgatga actgtgggtt ctgcatgaca gtatcactgg
 1601 tgtacagctc actctcgccc cggatgcggc ggaaccgagg gatcttccca
 1651 ttgtattcct gcttccggat ggtggaatac acctcgcccc cttcaaacag
 1701 aaccagggag ttctcgtccg ggctgaaggg ggctgagcct ctcatctcgc
 1751 caagtggcac cacagtgcc aaccaggt tccagcagct ggggtgccgg
 1801 gcgttgggtg cacaggccag cagccccca ctccgcctct ccaggagagt
 1851 gatgtagttc tcgcagtcct gcttatccag acaggacccc tttgtggagc
 1901 cgatattcac cgtgcgcaca gatgcgttct tgccctcggg gaagtcaaag
 1951 aggtagacct tgccacgtcc tcccaccac acagaggagc tgcctggctc
 2001 gtggaaaagc accgtgtgcg gctcagtctg gccaaagtcc acccggtcct
 2051 gccctacatg gcctttccag acggcgaaga tgcggggctc gctccttagg
 2101 tggccctggg cggaggcggc ggccgcccag agcagcagca gcagccgcag
 2151 ccgcagcgga agccccaacc gagccggcgg gccagggacg cgggcgcgcg
 2201 gtgcgctggg ggcggcacgt ccgggcggag gaggcgtcat cccaagccga
 2251 attcTGCAGA TATCCATCAC ACTGGCGGCC GCTCGAGCAT GCATCTAGAG
 2301 GGCCCAATTC GCCCTATAGT GAGTCGTATT ACAATTCAT GGCCGTCGTT
 2351 TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT

5 2401 TGCAGCACAT CCCCCTTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA
2451 CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG
2501 CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT
2551 GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC
2601 CTTCTTTTCT CGCCACGTTT GCCGGCTTTC CCCGTCAAGC TCTAAATCGG
10 2651 GGGCTCCCTT TAGGGTTCCG ATTTAGAGCT TTACGGCACC TCGACCGCAA
2701 AAAACTTGAT TTGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA
2751 CGGTTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCCTTAA TAGTGGACTC
2801 TTGTTCCAAA CTGGAACAAC ACTCAACCCT ATCGCGGTCT ATTCTTTTGA
2851 TTTATAAGGG ATTTTGCCGA TTTTCGCCCTA TTGGTTAAAA AATGAGCTGA
15 2901 TTTAACAAAT TCAGGGCGCA AGGGCTGCTA AAGGAACCGG AACACGTAGA
2951 AAGCCAGTCC GCAGAAACGG TGCTGACCCC GGATGAATGT CAGCTACTGG
3001 GCTATCTGGA CAAGGGAAAA CGCAAGCGCA AAGAGAAAGC AGGTAGCTTG
20 3051 CAGTGGGCTT ACATGGCGAT AGCTAGACTG GCGGTTTTTA TGGACAGCAA
3101 GCGAACCGGA ATTGCCAGCT GGGGCGCCCT CTGGTAAGGT TGGGAAGCCC
3151 TGCAAAGTAA ACTGGATGGC TTTCTTGCCG CCAAGGATCT GATGGCGCAG
3201 GGGATCAAGA TCTGATCAAG AGACAGGATG AGGATCGTTT CGCATGATTG
25 3251 AACAAGATGG ATTGACGCA GGTCTCCGG CCGCTTGGGT GGAGAGGCTA
3301 TTCGGCTATG ACTGGGCACA ACAGACAATC GGCTGCTCTG ATGCCGCCGT
3351 GTTCCGGCTG TCAGCGCAGG GCGCGCCGGT TCTTTTTGTC AAGACCGACC
3401 TGTCCGGTGC CCTGAATGAA CTGCAGGACG AGGCAGCGCG GCTATCGTGG
30 3451 CTGGCCACGA CGGGCGTTCC TTGCGCAGCT GTGCTCGACG TTGTCACTGA
3501 AGCGGGAAGG GACTGGCTGC TATTGGGCGA AGTGCCGGGG CAGGATCTCC
3551 TGTCATCTCG CCTTGCTCCT GCCGAGAAAG TATCCATCAT GGCTGATGCA
3601 ATGCGGCGGC TGCATACGCT TGATCCGGCT ACCTGCCCAT TCGACCACCA
35 3651 AGCGAAACAT CGCATCGAGC GAGCACGTAC TCGGATGGAA GCCGGTCTTG
3701 TCGATCAGGA TGATCTGGAC GAAGAGCATC AGGGGCTCGC GCCAGCCGAA
3751 CTGTTCGCCA GGCTCAAGGC GCGCATGCCC GACGGCGAGG ATCTCGTCGT
3801 GATCCATGGC GATGCCTGCT TGCCGAATAT CATGGTGGAA AATGGCCGCT
40 3851 TTTCTGGATT CAACGACTGT GGCCGGCTGG GTGTGGCGGA CCGCTATCAG
3901 GACATAGCGT TGGATACCCG TGATATTGCT GAAGAGCTTG GCGGCGAATG
3951 GGCTGACCGC TTCCTCGTGC TTTACGGTAT CGCCGCTCCC GATTCGCAGC
4001 GCATCGCCTT CTATCGCCTT CTTGACGAGT TCTTCTGAAT TGAAAAAGGA
45 4051 AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC CTTTTTTGCG
4101 GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA
4151 AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC
50 4201 TCAACAGCGG TAAGATCCTT GAGAGTTTTT GCCCGAAGA ACGTTTTCCA
4251 ATGATGAGCA CTTTTAAAGT TCTGCTATGT CATACACTAT TATCCCGTAT
4301 TGACGCCGGG CAAGAGCAAC TCGGTCGCCG GCGCGGTAT TCTCAGAATG

4351 ACTTGGTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC GGATGGCATG
 4401 ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC
 5 4451 GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT
 4501 TTTTGCACAA CATGGGGGAT CATGTAAGTC GCCTTGATCG TTGGGAACCG
 4551 GAGCTGAATG AAGCCATACC AAACGACGAG AGTGACACCA CGATGCCTGT
 4601 AGCAATGCCA ACAACGTTGC GCAAACTATT AACTGGCGAA CTACTTACTC
 10 4651 TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA
 4701 GGACCACTTC TGCCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA
 4751 ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA GCACTGGGGC
 4801 CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG
 15 4851 GCAACTATGG ATGAACGAAA TAGACAGATC GCTGAGATAG GTGCCTCACT
 4901 GATTAAGCAT TGGTAACTGT CAGACCAAGT TTACTCATAT ATACTTTAGA
 4951 TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT GAAGATCCTT
 5001 TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG
 20 5051 AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT
 5101 TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCACC GCTACCAGCG
 5151 GTGGTTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC
 5201 TGGCTTCAGC AGAGCGCAGA TACCAAATAC TGTCTTCTA GTGTAGCCGT
 25 5251 AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC ATACCTCGCT
 5301 CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT
 5351 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG
 5401 GCTGAACGGG GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC
 30 5451 ACCGAACTGA GATACCTACA GCGTGAGCAT TGAGAAAGCG CCACGCTTCC
 5501 CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG
 5551 GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA ACGCCTGGTA TCTTTATAGT
 35 5601 CCTGTCGGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT TGTGATGCTC
 5651 GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTAC
 5701 GGTTCCCTGGC CTTTTGCTGG CTTTTTGCTC ACATGTTCTT TCCTGCGTTA
 5751 TCCCCTGATT CTGTGGATAA CCGTATTACC GCCTTTGAGT GAGCTGATAC
 40 5801 CGCTCGCCGC AGCCGAACGA CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG
 5851 CGGAAG

45 Tabelle 8: Nukleotidsequenz des rekombinanten Expressionsplasmids pCDNA3.1(-)H-SemaL-MychHisA
 (SEQ ID NO.: 35)

50 1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC
 51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT
 101 GGAGGTCGCT GAGTAGTGCG CGAGCAAAAT TTAAGCTACA ACAAGGCAAG

151 GCTTGACCGA CAATTGCATG AAGAATCTGC TTAGGGTTAG GCGTTTTGCG
 201 CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT GATTATTGAC
 5 251 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG
 351 CCCAACGACC CCCGCCATT GACGTCAATA ATGACGTATG TTCCCATAGT
 401 AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAC TATTTACGGT
 10 451 AAAGTGGCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC
 501 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA
 551 CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
 601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
 15 651 TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA
 701 TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA
 751 ACAACTCCGC CCCATTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG
 801 GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA CTGCTTACTG
 20 851 GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
 901 GTTTAAACGG GCCCTCTAGA CTCGAGCGGC CGCCACTGTG CTGGATATCT
 951 GCAGaattcg gcttgggatg acgcctcctc cgcccggaag tgccgcccc
 1001 agcgacccgc gcgcccgagt ccctggcccg ccggctcggg tggggcttcc
 25 1051 gctgcggtg cggtgctgc tgcgtctctg ggcgcccgcc gcctccgccc
 1101 agggccacct aaggagcggg ccccgcatct tcgccgtctg gaaaggccat
 1151 gtagggcagg accgggtgga ctttggccag actgagccgc acacggtgct
 30 1201 tttccacgag ccaggcagct cctctgtgtg ggtgggagga cgtggcaagg
 1251 tctacctctt tgacttcccc gagggcaaga acgcatctgt gcgcacggtg
 1301 aatatcggt ccacaaagg gtcctgtctg gataagcggg actgcgagaa
 1351 ctacatcact ctccctggaga ggcgagtgga ggggctgctg gcctgtggca
 35 1401 ccaacgcccg gcaccccgag tgctggaacc tggatgaatg cactgtggtg
 1451 ccacttggtg agatgagagg ctacgcccc ttcagcccg acgagaactc
 1501 cctggttctg tttgaagggg acgaggtgta ttccaccatc cggaagcagg
 1551 aatacaatgg gaagatccct cggttccgcc gcatccgggg cgagagtgg
 40 1601 ctgtacacca gtgatactgt catgcagaac ccacagttca tcaaagccac
 1651 catcgtgcac caagaccagg cttacgatga caagatctac tacttcttcc
 1701 gagaggacaa tcctgacaag aatcctgagg ctccctctca tgtgtcccgt
 1751 gtggcccagt tgtgcagggg ggaccagggg ggggaaagt cactgtcagt
 45 1801 ctccaagtgg aacacttttc tgaaagccat gctggtatgc agtgatgctg
 1851 ccaccaacaa gaacttcaac aggtgcaag acgtcttct gtccttgac
 1901 ccagcggcc agtggaggga caccaggtc tatggtgtt tctccaaccc
 50 1951 ctggaactac tcagccgtct gtgtgtatc cctcggtgac attgacaagg
 2001 tcttccgtac ctccctcactc aagggtacc actcaagcct tcccaacccg
 2051 cggcctggca agtgccctcc agaccagcag ccgataccca cagagacctt

5

10

15

20

25

30

35

40

45

50

55

2101 ccaggtggct gaccgtcacc cagaggtggc gcagaggggtg gagcccatgg
 2151 ggcctctgaa gacgccattg ttccactcta aataccacta ccagaaagtg
 2201 gccgttcacc gcatgcaagc cagccacggg gagacctttc atgtgcttta
 2251 cctaactaca gacaggggca ctatccacaa ggtggtggaa ccgggggagc
 2301 aggagcacag cttcgccttc aacatcatgg agatccagcc ctccgcgcgc
 2351 gcggctgcca tccagaccat gtcgctggat gctgagcgga ggaagctgta
 2401 tgtgagctcc cagtgggagg tgagccaggt gcccctggac ctgtgtgagg
 2451 tctatggcgg gggctgccac ggttgccctca tgtcccgaga cccctactgc
 2501 ggctgggacc agggccgctg catctccatc tacagctccg aacggtcagt
 2551 gctgcaatcc attaateccag ccgagccaca caaggagtgt cccaacccca
 2601 aaccagacaa ggccccactg cagaaggttt ccctggcccc aaactctcgc
 2651 tactacctga gctgccccat ggaatcccg caccgccact actcatggcg
 2701 ccacaaggag aacgtggagc agagctgcga acctggtcac cagagcccca
 2751 actgcatcct gtccatcgag aacctcacgg cgcagcagta cggccactac
 2801 ttctgcgagg cccaggaggg ctctacttc cgcgaggctc agcactggca
 2851 gctgctgccc gaggacggca tcatggccga gcacctgctg ggtcatgcct
 2901 gtgccctggc tgcctccctc tggctggggg tgctgcccac actcactctt
 2951 ggcttgctgg tccacgtgaa gcttGGGCCC GAACAAAAAC TCATCTCAGA
 3001 AGAGGATCTG AATAGCGCCG TCGACCATCA TCATCATCAT CATTGAGTTT
 3051 AAACCGCTGA TCAGCCTCGA CTGTGCCTTC TAGTTGCCAG CCATCTGTTG
 3101 TTTGCCCTC CCCCCTGCCT TCCTTGACCC TGGAAAGGTGC CACTCCCCT
 3151 GTCCTTCTCT AATAAAATGA GGAAATTGCA TCGCATTGTC TGAGTAGGTG
 3201 TCATTCTATT CTGGGGGGTG GGGTGGGGCA GGACAGCAAG GGGGAGGATT
 3251 GGGAAGACAA TAGCAGGCAT GCTGGGGATG CGGTGGGCTC TATGGCTTCT
 3301 GAGGCGGAAA GAACCAGCTG GGGCTCTAGG GGGTATCCCC ACGCGCCCTG
 3351 TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT GGTTACGCGC AGCGTGACCG
 3401 CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC
 3451 TTTCTCGCCA CGTTCGCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGCAT
 3501 CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCAAAAAAAC
 3551 TTGATTAGGG TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT
 3601 TTTGCCCCTT TGACGTTGGA GTCCACGTTT TTTAATAGTG GACTCTTGTT
 3651 CCAAACTGGA ACAACACTCA ACCCTATCTC GGTCTATTCT TTTGATTAT
 3701 AAGGGATTTT GGGGATTTTC GCCTATTGGT TAAAAATGA GCTGATTTAA
 3751 CAAAAATTTA ACGCGAATTA ATTCTGTGGA ATGTGTGTCA GTTAGGGTGT
 3801 GGAAAGTCCC CAGGCTCCCC AGGCAGGCAG AAGTATGCAA AGCATGCATC
 3851 TCAATTAGTC AGCAACCAGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC
 3901 AGAAGTATGC AAAGCATGCA TCTCAATTAG TCAGCAACCA TAGTCCCGCC
 3951 CCTAACTCCG CCCATCCCGC CCCTAACTCC GCCCAGTTCC GCCCATCTCT
 4001 CGCCCCATGG CTGACTAATT TTTTATTAT ATGCAGAGGC CGAGGCCGCC

4051 TCTGCCTCTG AGCTATTCCA GAAGTAGTGA GGAGGCTTTT TTGGAGGCCT
4101 AGGCTTTTGC AAAAAGCTCC CGGGAGCTTG TATATCCATT TTCGGATCTG
5 4151 ATCAAGAGAC AGGATGAGGA TCGTTTCGCA TGATTGAACA AGATGGATTG
4201 CACGCAGGTT CTCCGGCCGC TTGGGTGGAG AGGCTATTCTG GCTATGACTG
4251 GGCACAACAG ACAATCGGCT GCTCTGATGC CGCCGTGTTC CGGCTGTCAG
4301 CGCAGGGGCG CCCGGTTCTT TTTGTCAAGA CCGACCTGTC CGGTGCCCTG
10 4351 AATGAAGTGC AGGACGAGGC AGCGCGGCTA TCGTGGCTGG CCACGACGGG
4401 CGTTCCTTGC GCAGCTGTGC TCGACGTTGT CACTGAAGCG GGAAGGGACT
4451 GGCTGCTATT GGGCGAAGTG CCGGGGCAGG ATCTCCTGTC ATCTCACCTT
4501 GCTCCTGCCG AGAAAGTATC CATCATGGCT GATGCAATGC GGCGGCTGCA
15 4551 TACGCTTGAT CCGGCTACCT GCCCATTCTGA CCACCAAGCG AAACATCGCA
4601 TCGAGCGAGC ACGTACTCGG ATGGAAGCCG GTCTTGTCGA TCAGGATGAT
4651 CTGGACGAAG AGCATCAGGG GCTCGCGCCA GCCGAAGTGT TCGCCAGGCT
4701 CAAGGCGCGC ATGCCCCGACG GCGAGGATCT CGTCGTGACC CATGGCGATG
20 4751 CCTGCTTGCC GAATATCATG GTGGAAAATG GCCGCTTTTC TGGATTTCATC
4801 GACTGTGGCC GGCTGGGTGT GCGGGACCGC TATCAGGACA TAGCGTTGGC
4851 TACCCGTGAT ATTGCTGAAG AGCTTGCGCG CGAATGGGCT GACCGCTTCC
4901 TCGTGCTTTA CCGTATCGCC GCTCCCGATT CGCAGCGCAT CGCCTTCTAT
25 4951 CGCCTTCTTG ACGAGTTCTT CTGAGCGGGA CTCTGGGGTT CGAAATGACC
5001 GACCAAGCGA CGCCCCAACC GCCATCACGA GATTTTCGATT CCACCGCCGC
5051 CTTCTATGAA AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA
30 5101 TGATCCTCCA GCGCGGGGAT CTCATGCTGG AGTTCCTTCG CCACCCCAAC
5151 TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA
5201 TTTCACAAAT AAAGCATTTT TTTCAGTCA TTCTAGTTGT GGTTTGTCCA
5251 AACTCATCAA TGTATCTTAT CATGTCTGTA TACCGTCGAC CTCTAGCTAG
35 5301 AGCTTGCGCT AATCATGGTC ATAGCTGTTT CCTGTGTGAA ATTGTTATCC
5351 GCTCACAATT CCACACAACA TACGAGCCGG AAGCATAAAG TGTAAGCCT
5401 GGGGTGCCTA ATGAGTGAGC TAACTCACAT TAATTGCGTT GCGCTCACTG
5451 CCCGCTTTCC AGTCGGGAAA CCTGTCTGTC CAGCTGCATT AATGAATCGG
40 5501 CCAACGCGCG GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT
5551 CGCTCACTGA CTCGCTGCGC TCGGTCTGTC GGCTGCGGCG AGCGGTATCA
5601 GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG GGGATAACGC
5651 AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCCTAAAA
45 5701 AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCC TGACGAGCAT
5751 CACAAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA
5801 AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC
5851 CGACCTGCC GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC
50 5901 GTGGCGCTTT CTCAATGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT
5951 CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCCGACC

5
10
15
20
25
30
35
40

```

6001 GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC
6051 GACTTATCGC CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG
6101 GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT
6151 ACACTAGAAG GACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC
6201 TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG
6251 TAGCGGTGGT TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
6301 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG
6351 AACGAAAAC TACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT
6401 CTTACCTAG ATCCTTTTAA ATTAATAATG AAGTTTAAA TCAATCTAAA
6451 GTATATATGA GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG
6501 GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG TTGCCTGACT
6551 CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGCCCCCA
6601 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTATCA
6651 GCAATAAACC AGCCAGCCGG AAGGGCCGAG CGCAGAAAGT GTCCTGCAAC
6701 TTTATCCGCC TCCATCCAGT CTATTAATTG TTGCCGGGAA GCTAGAGTAA
6751 GTAGTTCCGC AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC
6801 ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC
6851 CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
6901 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAAGT
6951 TTATCACTCA TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC
7001 ATCCGTAAGA TGCTTTTCTG TGAAGTGGTGA GTACTCAACC AAGTCATTCT
7051 GAGAATAGTG TATGCGGCGA CCGAGTTGCT CTTGCCCGGC GTCAATACGG
7101 GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGAAAA
7151 ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
7201 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT
7251 TTCACCAGCG TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA
7301 AAAGGGAATA AGGGCGACAC GGAAATGTTG AATACTCATA CTCTTCCTTT
7351 TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC
7401 ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT CGCGCACATT
7451 TCCCCGAAAA GTGCCACCTG ACGTC

```

45
Tabelle 9: Nukleotidsequenz des rekombinanten Plasmids pcDNA 3.1-H-SemaL-EGFP-MychisA
(SEQ ID NO.: 36)

50
55

```

1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC
51 TGCTCTGATG CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT
101 GGAGGTCGCT GAGTAGTGCG CGAGCAAAAT TTAAGCTACA ACAAGGCAAG
151 GCTTGACCGA CAATTGCATG AAGAATCTGC TTAGGGTTAG GCGTTTTGCG
201 CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT GATTATTGAC

```

251 TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG
 5 351 CCCAACGACC CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT
 401 AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAC TATTTACGGT
 451 AAAGTGGCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTACGCCC
 501 CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA
 10 551 CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
 601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA
 651 TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA
 701 TGGGAGTTTG TTTTGGCACC AAAATCAACG GGAAGTTTCCA AAATGTCGTA
 15 751 ACAACTCCGC CCCATTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG
 801 GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA CTGCTTACTG
 851 GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
 901 GTTTAAACGG GCCCTCTAGA CTCGAGCGGC CGCCACTGTG CTGGATATCT
 20 951 GCAGaattcg gcttgggatg acgcctcctc cgcccgagcg tgccgcccc
 1001 agcgacccgc gcgcccgcgt ccttggcccg ccggtcgggt tggggcttcc
 1051 gctgcggctg cggtcgtctg tgctgctctg ggcgcccgcc gcctccgccc
 1101 agggccacct aaggagcgga ccccgcatct tcgccgtctg gaaaggccat
 25 1151 gtagggcagg accgggtgga ctttggccag actgagccgc acacggtgct
 1201 tttccacgag ccaggcagct cctctgtgtg ggtgggagga cgtggcaagg
 1251 tctacctctt tgacttcccc gagggcaaga acgcatctgt gcgcacggtg
 1301 aatatcggtt ccacaaaggg gtccgtgtctg gataagcggg actgagagaa
 1351 ctacatcact ctccctggaga ggcgagtgga ggggctgctg gcctgtggca
 1401 ccaacgcccg gcaccccagc tgctggaacc tgggtgaatg cactgtggtg
 1451 ccacttggcg agatgagagg ctacgcccc ttcagcccgg acgagaactc
 35 1501 cctggttctg tttgaagggg acgaggtgta ttccaccatc cggaagcagg
 1551 aatacaatgg gaagatccct cggttccgcc gcatccgggg cgagagttag
 1601 ctgtacacca gtgatactgt catgcagaac ccacagttca tcaaagccac
 1651 catcgtgcac caagaccagg cttacgatga caagatctac tacttcttcc
 40 1701 gagaggacaa tcctgacaag aatcctgagg ctccctctca tgtgtcccgt
 1751 gtggcccagt tgtgcagggg ggaccagggt ggggaaagt cactgtcagt
 1801 ctccaagtgg aacacttttc tgaaagccat gctggtatgc agtgatgctg
 1851 ccaccaacaa gaacttcaac aggctgcaag acgtcttcc tctccctgac
 45 1901 ccagcgggc agtggaggga caccagggtc tatggtgtt tctccaaccc
 1951 ctggaactac tcagccgtct gtgtgtattc cctcggtgac attgacaagg
 2001 tcttccgtac ctccctcactc aagggtacc actcaagcct tcccaacccg
 50 2051 cggcctggca agtgccctccc agaccagcag ccgataccca cagagacctt
 2101 ccaggtggct gaccgtcac cagaggtggc gcagaggggt gagcccatgg
 2151 ggcctctgaa gacgccattg ttccactcta aataccacta ccagaaagtg

5

10

15

20

25

30

35

40

45

50

55

2201 gccgttcacc gcatgcaagc cagccacggg gagacctttc atgtgcttta
 2251 cctaactaca gacaggggca ctatccacaa ggtggtggaa ccgggggagc
 2301 aggagcacag cttcgccttc aacatcatgg agatccagcc cttccgccgc
 2351 gcggctgcca tccagaccat gtcgctggat gctgagcgga ggaagctgta
 2401 tgtgagctcc cagtgggagg tgagccaggt gcccctggac ctgtgtgagg
 2451 tctatggcgg gggctgccac ggttgccctca tgtcccgaga cccctactgc
 2501 ggctgggacc agggccgctg catctccatc tacagctccg aacggtcagt
 2551 gctgcaatcc attaatccag ccgagccaca caaggagtgt cccaacccca
 2601 aaccagacaa ggccccactg cagaaggttt ccctggcccc aaactctcgc
 2651 tactacctga gctgccccat ggaatcccgc cagccacct actcatggcg
 2701 ccacaaggag aacgtggagc agagctgcga acctggtcac cagagcccca
 2751 actgcatcct gttcatcgag aacctcacgg cgcagcagta cggccactac
 2801 ttctgcgagg cccaggaggg ctctacttc cgcgaggctc agcactggca
 2851 gctgctgccc gaggacggca tcatggccga gcacctgctg ggtcatgcct
 2901 gtgccctggc tgccctccctc tggctggggg tgctgcccac actcactctt
 2951 ggcttgctgg tccacATGGT GAGCAAGGGC GAGGAGCTGT TCACCGGGGT
 3001 GGTGCCCATC CTGGTCGAGC TGGACGGCGA CGTAAACGGC CACAAGTTCA
 3051 GCGTGTCCGG CGAGGGCGAG GGCGATGCCA CCTACGGCAA GCTGACCCTG
 3101 AAGTTCATCT GCACCACCGG CAAGCTGCCC GTGCCCTGGC CCACCCTCGT
 3151 GACCACCCTG ACCTACGGCG TGCAGTGCTT CAGCCGCTAC CCCGACCACA
 3201 TGAAGCAGCA CGACTTCTTC AAGTCCGCCA TGCCCCGAAGG CTACGTCCAG
 3251 GAGCGCACCA TCTTCTTCAA GGACGACGGC AACTACAAGA CCCGCGCCGA
 3301 GGTGAAGTTC GAGGGCGACA CCCTGGTGAA CCGCATCGAG CTGAAGGGCA
 3351 TCGACTTCAA GGAGGACGGC AACATCCTGG GGCACAAGCT GGAGTACAAC
 3401 TACAACAGCC ACAACGTCTA TATCATGGCC GACAAGCAGA AGAACGGCAT
 3451 CAAGGTGAAC TTCAAGATCC GCCACAACAT CGAGGACGGC AGCGTGCAGC
 3501 TCGCCGACCA CTACCAGCAG AACACCCCCA TCGGCGACGG CCCCCTGCTG
 3551 CTGCCCCGACA ACCACTACCT GAGCACCCAG TCCGCCCTGA GCAAAGACCC
 3601 CAACGAGAAG CGCGATCACA TGGTCCTGCT GGAGTTCGTG ACCGCCGCCG
 3651 GGATCACTCT CGGCATGGAC GAGCTGTACA Aggtgaagct tGGGCCCCGAA
 3701 CAAAAACTCA TCTCAGAAGA GGATCTGAAT AGCGCCGTCG ACCATCATCA
 3751 TCATCATCAT TGAGTTTAAA CCGCTGATCA GCCTCGACTG TGCCTTCTAG
 3801 TTGCCAGCCA TCTGTTGTTT GCCCTCCCC CGTGCCTTCC TTGACCCTGG
 3851 AAGGTGCCAC TCCCCTGTCT CTTTCTTAAT AAAATGAGGA AATTGCATCG
 3901 CATTGTCTGA GTAGGTGTCA TTCTATTCTG GGGGGTGGGG TGGGGCAGGA
 3951 CAGCAAGGGG GAGGATTGGG AAGACAATAG CAGGCATGCT GGGGATGCGG
 4001 TGGGCTCTAT GGCTTCTGAG GCGGAAAGAA CCAGCTGGGG CTCTAGGGGG
 4051 TATCCCCACG CGCCCTGTAG CGGCGCATTG AGCGCGGCGG GTGTGGTGGT
 4101 TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT

5 4151 TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA
4201 GCTCTAAATC GGGGCATCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA
4251 CCTCGACCCC AAAAACTTG ATTAGGGTGA TGGTTCACGT AGTGGGCCAT
4301 CGCCCTGATA GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT
4351 AATAGTGGAC TCTTGTCCA AACTGGAACA ACACTCAACC CTATCTCGGT
4401 CTATTCTTTT GATTTATAAG GGATTTTGGG GATTTTCGGC TATTGGTTAA
10 4451 AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTAATT CTGTGGAATG
4501 TGTGTCAGTT AGGGTGTGGA AAGTCCCCAG GCTCCCCAGG CAGGCAGAAG
4551 TATGCAAAGC ATGCATCTCA ATTAGTCAGC AACCAGGTGT GGAAAGTCCC
4601 CAGGCTCCCC AGCAGGCAGA AGTATGCAAA GCATGCATCT CAATTAGTCA
15 4651 GCAACCATAG TCCCGCCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC
4701 CAGTTCCGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG
4751 CAGAGGCCGA GGCCGCCTCT GCCTCTGAGC TATTCCAGAA GTAGTGAGGA
4801 GGCTTTTTTG GAGGCCTAGG CTTTTCGAAA AAGCTCCCGG GAGCTTGAT
20 4851 ATCCATTTTC GGATCTGATC AAGAGACAGG ATGAGGATCG TTTCCGATGA
4901 TTGAACAAGA TGGATTGCAC GCAGTTCTC CGGCCGCTTG GGTGGAGAGG
4951 CTATTCGGCT ATGACTGGGC ACAACAGACA ATCGGCTGCT CTGATGCCGC
5001 CGTGTTCGG CTGTCAGCGC AGGGGCGCCC GGTTCTTTTT GTCAAGACCG
25 5051 ACCTGTCCGG TGCCCTGAAT GAACTGCAGG ACGAGGCAGC GCGGCTATCG
5101 TGGCTGGCCA CGACGGGCGT TCCTTGCGCA GCTGTGCTCG ACGTTGTCAC
5151 TGAAGCGGGA AGGGACTGGC TGCTATTGGG CGAAGTGCCG GGGCAGGATC
30 5201 TCCTGTCATC TCACCTTGCT CCTGCCGAGA AAGTATCCAT CATGGCTGAT
5251 GCAATGCGGC GGCTGCATAC GCTTGATCCG GCTACCTGCC CATTGCACCA
5301 CCAAGCGAAA CATCGCATCG AGCGAGCAGC TACTCGGATG GAAGCCGGTC
5351 TTGTCGATCA GGATGATCTG GACGAAGAGC ATCAGGGGCT CGCGCCAGCC
35 5401 GAACTGTTCG CCAGGCTCAA GGCGCGCATG CCCGACGGCG AGGATCTCGT
5451 CGTGACCCAT GGCGATGCCT GCTTGCCGAA TATCATGGTG GAAAATGGCC
5501 GCTTTTCTGG ATTCATCGAC TGTGGCCGGC TGGGTGTGGC GGACCGCTAT
5551 CAGGACATAG CGTTGGCTAC CCGTGATATT GCTGAAGAGC TTGGCGGCGA
40 5601 ATGGGCTGAC CGCTTCCTCG TGCTTTACGG TATCGCCGCT CCCGATTTCG
5651 AGCGCATCGC CTTCTATCGC CTTCTTGACG AGTTCTTCTG AGCGGGACTC
5701 TGGGGTTCGA AATGACCGAC CAAGCGACGC CCAACCTGCC ATCAGAGAT
5751 TTCGATTCCA CCGCCGCCTT CTATGAAAGG TTGGGCTTCG GAATCGTTTT
45 5801 CCGGGACGCC GGCTGGATGA TCCTCCAGCG CGGGGATCTC ATGCTGGAGT
5851 TCTTCGCCCC CCCCAACTTG TTTATTGCAG CTTATAATGG TTACAAATAA
5901 AGCAATAGCA TCACAAATTT CACAAATAAA GCATTTTTTT CACTGCATT
5951 TAGTTGTGGT TTGTCCAAAC TCATCAATGT ATCTTATCAT GTCTGTATAC
50 6001 CGTCGACCTC TAGCTAGAGC TTGGCGTAAT CATGGTCATA GCTGTTTCCT
6051 GTGTGAAATT GTTATCCGCT CACAATTCCA CACAACATAC GAGCCGGAAG

5

10

15

20

25

30

35

40

45

50

55

6101 CATAAAGTGT AAAGCCTGGG GTGCCTAATG AGTGAGCTAA CTCACATTAA
 6151 TTGCGTTGCG CTCACTGCCC GCTTTCCAGT CGGGAAACCT GTCGTGCCAG
 6201 CTGCATTAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT TGC GTATTGG
 6251 GCGCTCTTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG GTCGTTCCGC
 6301 TGCGGCGAGC GGTATCAGCT CACTCAAAGG CGGTAATACG GTTATCCACA
 6351 GAATCAGGGG ATAACGCAGG AAAGAACATG TGAGCAAAAG GCCAGCAAAA
 6401 GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GGCGTTTTTC CATAGGCTCC
 6451 GCCCCCTGA CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA
 6501 AACCCGACAG GACTATAAAG ATACCAGGCG TTTCCCCCTG GAAGCTCCCT
 6551 CGTGCGCTCT CCTGTTCCGA CCCTGCCGCT TACCGGATAC CTGTCCGCCT
 6601 TTCTCCCTTC GGGAAAGCGT GCGCTTTCTC AATGCTCACG CTGTAGGTAT
 6651 CTCAGTTCGG TGTAGGTCGT TCGCTCCAAG CTGGGCTGTG TGCACGAACC
 6701 CCCC GTTCAG CCGACCGCT GCGCCTTATC CGGTA ACTAT CGTCTTGAGT
 6751 CCAACCCGGT AAGACACGAC TTATCGCCAC TGGCAGCAGC CACTGGTAAC
 6801 AGGATTAGCA GAGCGAGGTA TG TAGGCGGT GCTACAGAGT TCTTGAAGTG
 6851 GTGGCCTAAC TACGGCTACA CTAGAAGGAC AGTATTGGT ATCTGCGCTC
 6901 TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC TTGATCCGGC
 6951 AAACAAACCA CCGCTGGTAG CGGTGGTTTT TTTGTTTGCA AGCAGCAGAT
 7001 TACGCGCAGA AAAAAAGGAT CTCAAGAAGA TCCTTTGATC TTTTCTACGG
 7051 GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG
 7101 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAATT AAAAATGAAG
 7151 TTTTAAATCA ATCTAAAGTA TATATGAGTA AACTTGGTCT GACAGTTACC
 7201 AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTGTTCA
 7251 TCCATAGTTG CCTGACTCCC CGTCGTGTAG ATA ACTACGA TACGGGAGGG
 7301 CTTACCATCT GGCCCCAGTG CTGCAATGAT ACCGCGAGAC CCACGCTCAC
 7351 CGGCTCCAGA TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCCGAGCGC
 7401 AGAAGTGGTC CTGCAACTTT ATCCGCCTCC ATCCAGTCTA TTAATTGTTG
 7451 CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG
 7501 TTGCCATTGC TACAGGCATC GTGGTGTCAC GTCGTCGTT TGGTATGGCT
 7551 TCATT CAGCT CCGGTTCCCA ACGATCAAGG CGAGTTACAT GATCCCCCAT
 7601 GTTGTGCAAA AAAGCGGTTA GTCCTTCGG TCCTCCGATC GTTGTCAGAA
 7651 GTAAGTTGGC CGCAGTGTTA TCACTCATGG TTATGGCAGC ACTGCATAAT
 7701 TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA CTGGTGAGTA
 7751 CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGCGACCG AGTTGCTCTT
 7801 GCCCCGGCGTC AATACGGGAT AATACCGCGC CACATAGCAG AACTTTAAAA
 7851 GTGCTCATCA TTGAAAACG TTCTTCGGGG CGAAACTCT CAAGGATCTT
 7901 ACCGCTGTTG AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT
 7951 CTT CAGCATC TTTTACTTTC ACCAGCGTTT CTGGGTGAGC AAAAACAGGA
 8001 AGGCAAAATG CCGCAAAAAA GGGAATAAGG GCGACACGGA AATGTTGAAT

8051 ACTCATACTC TTCCTTTTTC AATATTATTG AAGCATTAT CAGGGTTATT
 8101 GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA TAAACAAATA
 8151 GGGGTTCCGC GCACATTTC CCGAAAAGTG CCACCTGACG TC

Tabelle 10: Nukleotidsequenz des rekombinanten Plasmids pIND-H-SemaL-EE (SEQ ID NO.: 37)

1 AGATCTCGGC CGCATATTAA GTGCATTGTT CTCGATACCG CTAAGTGCAT
 51 TGTTCTCGTT AGCTCGATGG ACAAGTGCAT TGTTCTCTTG CTGAAAGCTC
 101 GATGGACAAG TGCATTGTTT TCTTGCTGAA AGCTCGATGG ACAAGTGCAT
 151 TGTTCTCTTG CTGAAAGCTC AGTACCCGGG AGTACCCTCG ACCGCCGGAG
 201 TATAAATAGA GGCCTTCGT CTACGGAGCG ACAATTCAAT TCAAACAAGC
 251 AAAGTGAACA CGTCGCTAAG CGAAAGCTAA GCAAATAAAC AAGCGCAGCT
 301 GAACAAGCTA AACAATCTGC AGTAAAGTGC AAGTTAAAGT GAATCAATTA
 351 AAAGTAACCA GCAACCAAGT AAATCAACTG CAACTACTGA AATCTGCCAA
 401 GAAGTAATTA TTGAATACAA GAAGAGAACT CTGAATACTT TCAACAAGTT
 451 ACCGAGAAAG AAGAACTCAC ACACAGCTAG CGTTTAACT TAAGCTTGGT
 501 ACCGAGCTCG GATCCACTAG TCCAGTGTGG TGgaattcgg cttgggatga
 551 cgctctctcc gcccggaagt gccgccccca gcgcaccgcg cgcccgcgctc
 601 cctggcccgcc cggctcgggtt ggggcttccg ctgcggctgc ggctgctgct
 651 gctgctctgg gcggccgccc cctccgcccc gggccacctt aggagcggac
 701 cccgcatctt cgccgtctgg aaaggccatg tagggcagga cggggtggac
 751 tttggccaga ctgagccgca cacggtgctt ttccacgagc caggcagctc
 801 ctctgtgtgg gtgggaggac gtggcaaggc ctacctctt gacttccccg
 851 agggcaagaa cgcatctgtg cgcacggtga atatcggctc cacaagggg
 901 tcctgtctgg ataagcggga ctgcgagAAC tacatcactc tcctggagag
 951 gcggagttag gggctgctgg cctgtggcac caacgcccgg caccagct
 1001 gctggaacct ggtgaatggc actgtggtgc cacttggcga gatgagaggc
 1051 tacgccccct tcagcccgga cgagaactcc ctggttctgt ttgaagggga
 1101 cgagggtgat tccaccatcc ggaagcagga atacaatggg aagatccctc
 1151 gggtccgccc catccggggc gagagtgagc tgtacaccag tgatactgtc
 1201 atgcagaacc cacagttcat caaagccacc atcgtgcacc aagaccaggc
 1251 ttacgatgac aagatctact acttcttccg agaggacaat cctgacaaga
 1301 atcctgaggc tcctctcaat gtgtcccgtg tggcccagtt gtgcagggg
 1351 gaccaggggtg gggaaagtcc actgtcagtc tccaagtgga acacttttct
 1401 gaaagccatg ctggtatgca gtgatgctgc caccaacaag aacttcaaca
 1451 ggctgcaaga cgtcttctctg ctccctgacc ccagcggcca gtggagggac
 1501 accagggctc atgggtgttt ctccaacccc tggaactact cagccgtctg
 1551 tgtgtattcc ctcggtgaca ttgacaaggc cttccgtacc tcctcactca

5

10

15

20

25

30

35

40

45

50

55

```

1601 aggggtacca ctcaagcctt cccaaccgc ggcttgcaa gtgcctccca
1651 gaccagcagc cgataccac agagacctc cagggtgctg accgtcacc
1701 agaggtggcg cagaggtgg agcccatgg gcctctgaag acgccattgt
1751 tccactctaa ataccactac cagaaagtgg ccgttcaccg catgcaagcc
1801 agccacgggg agacctttca tgtgctttac ctaactacag acaggggcac
1851 tatccacaag gtggtggaac cgggggagca ggagcacagc ttcgccttca
1901 acatcatgga gatccagccc ttccgccgcg cggctgccat ccagaccatg
1951 tcgctggatg ctgagcggag gaagctgtat gtgagctccc agtgggaggt
2001 gagccaggtg cccctggacc tgtgtgaggt ctatggcggg ggctgccacg
2051 gttgcctcat gtcccagac ccctactgcg gctgggacca gggccgctgc
2101 atctccatct acagctccga acggtcagtg ctgcaatcca ttaatccagc
2151 cgagccacac aaggagtgtc ccaaccccaa accagacaag gccccactgc
2201 agaaggtttc cctggcccca aactctcgct actacctgag ctgccccatg
2251 gaatcccgcg acgccaccta ctcatggcgc cacaaggaga acgtggagca
2301 gagctgcgaa cctggtcacc agagcccaa ctgcatcctg ttcatcgaga
2351 acctcacggc gcagcagtac ggccactact tctgcgaggc ccaggagggc
2401 tctactttcc gcgaggctca gcaactggcag ctgctgcccg aggacggcat
2451 catggccgag cacctgctgg gtcattgcctg tgccctggct gcctccctct
2501 ggetgggggt gctgccaca ctcaactctg gcttgcctgt ccacgtgaag
2551 cttGGGCCCC TTTAAACCCG CTGATCAGCC TCGACTGTGC CTTCTAGTTG
2601 CCAGCCATCT GTTGTGTTGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG
2651 GTGCCACTCC CACTGTCCTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT
2701 TGTCTGAGTA GGTGTCATTC TATTCTGGGG GGTGGGGTGG GGCAGGACAG
2751 CAAGGGGGAG GATTGGGAAG ACAATAGCAG GCATGCTGGG GATGCGGTGG
2801 GCTCTATGGC TTCTGAGGCG GAAAGAACCA GCTGGGGCTC TAGGGGGTAT
2851 CCCCACGCGC CCTGTAGCGG CGCATTAAGC GCGGCGGGTG TGGTGGTTAC
2901 GCGCAGCGTG ACCGCTACAC TTGCCAGCGC CCTAGCGCCC GCTCCTTTCG
2951 CTTTCTTCCC TTCCTTTC TCACGTTTCG CCGGCTTTC CCGTCAAGCT
3001 CTAAATCGGG GCATCCCTTT AGGGTTCCGA TTTAGTGCTT TACGGCACCT
3051 CGACCCCAAA AACTTGATT AGGGTGATGG TTCACGTAGT GGGCCATCGC
3101 CCTGATAGAC GGTTTTTCGC CCTTTGACGT TGGAGTCCAC GTTCTTTAAT
3151 AGTGGACTCT TGTTCCAAAC TGAACAACA CTCAACCCTA TCTCGGTCTA
3201 TTCTTTTGAT TTATAAGGGA TTTTGGGGAT TTCGGCCTAT TGGTTAAAAA
3251 ATGAGCTGAT TTAACAAAAA TTTAACGCGA ATTAATTCTG TGAATGTGT
3301 GTCAGTTAGG GTGTGAAAG TCCCCAGGCT CCCCAGGCAG GCAGAAGTAT
3351 GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGTGTGGA AAGTCCCCAG
3401 GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA
3451 ACCATAGTCC CGCCCCTAAC TCCGCCATC CCGCCCCTAA CTCCGCCAG
3501 TTCCGCCCAT TCTCCGCCCC ATGGCTGACT AATTTTTTTT ATTTATGCAG

```

3551 AGGCCGAGGC CGCCTCTGCC TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC
 3601 TTTTTTGGAG GCCTAGGCTT TTGCAAAAAG CTCCCGGGAG CTTGTATATC
 5 3651 CATTTCGGA TCTGATCAAG AGACAGGATG AGGATCGTTT CGCATGATTG
 3701 AACAAGATGG ATTGCACGCA GGTTCCTCCGG CCGCTTGGGT GGAGAGGCTA
 3751 TTCGGCTATG ACTGGGCACA ACAGACAATC GGCTGCTCTG ATGCCGCCGT
 3801 GTTCCGGCTG TCAGCGCAGG GGCGCCCGGT TCTTTTGTG AAGACCGACC
 10 3851 TGTCCGGTGC CCTGAATGAA CTGCAGGACG AGGCAGCGCG GCTATCGTGG
 3901 CTGGCCACGA CGGGCGTTCC TTGCGCAGCT GTGCTCGACG TTGTACTGA
 3951 AGCGGGAAGG GACTGGCTGC TATTGGGCGA AGTGCCGGGG CAGGATCTCC
 4001 TGTCATCTCA CCTTGCTCCT GCCGAGAAAG TATCCATCAT GGCTGATGCA
 15 4051 ATGCGGCGGC TGCATACGCT TGATCCGGCT ACCTGCCCAT TCGACCACCA
 4101 AGCGAAACAT CGCATCGAGC GAGCACGTAC TCGGATGGAA GCCGGTCTTG
 4151 TCGATCAGGA TGATCTGGAC GAAGAGCATC AGGGGCTCGC GCCAGCCGAA
 4201 CTGTTTCGCCA GGCTCAAGGC GCGCATGCCC GACGGCGAGG ATCTCGTCGT
 20 4251 GACCCATGGC GATGCCTGCT TGCCGAATAT CATGGTGGAA AATGGCCGCT
 4301 TTTCTGGATT CATCGACTGT GGCCGGCTGG GTGTGGCGGA CCGCTATCAG
 4351 GACATAGCGT TGGCTACCCG TGATATTGCT GAAGAGCTTG GCGGCGAATG
 4401 GGCTGACCGC TTCCTCGTGC TTTACGGTAT CGCCGCTCCC GATTGCGAGC
 25 4451 GCATCGCCTT CTATCGCCTT CTTGACGAGT TCTTCTGAGC GGGACTCTGG
 4501 GGTTTCGAAAT GACCGACCAA GCGACGCCCC ACCTGCCATC ACGAGATTTT
 4551 GATTCCACCG CCGCCTTCTA TGAAAGGTTG GGCTTCGGAA TCGTTTTCCG
 30 4601 GGACGCCGGC TGGATGATCC TCCAGCGCGG GGATCTCATG CTGGAGTTCT
 4651 TCGCCCACCC CAACTTGTTT ATTGCAGCTT ATAATGGTTA CAAATAAAGC
 4701 AATAGCATCA CAAATTTTAC AAATAAAGCA TTTTTTTCAC TGCATTCTAG
 4751 TTGTGGTTTG TCCAAACTCA TCAATGTATC TTATCATGTC TGTATACCGT
 35 4801 CGACCTCTAG CTAGAGCTTG GCGTAATCAT GGTCATAGCT GTTTCCTGTG
 4851 TGAAATTGTT ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT
 4901 AAAGTGTAAG GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG
 4951 CGTTGCGCTC ACTGCCCCGT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG
 40 5001 CATTAATGAA TCGGCCAACG CGCGGGGAGA GGCGGTTTGC GTATTGGGCG
 5051 CTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC
 5101 GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA
 5151 TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC
 45 5201 CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCAT AGGCTCCGCC
 5251 CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC
 5301 CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT
 5351 GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCCTTTC
 50 5401 TCCCTTCGGG AAGCGTGGCG CTTTCTCAAT GCTCACGCTG TAGGTATCTC
 5451 AGTTCGGTGT AGGTCGTTTC CTCCAAGCTG GGCTGTGTGC ACGAACCCCC

5501 CGTTCAGCCC GACCGCTGCG CTTATCCGG TAACTATCGT CTTGAGTCCA
 5551 ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG
 5601 ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG
 5651 GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TGCCTCTGCG
 5701 TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA
 5751 CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC
 5801 GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT
 5851 CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA
 5901 TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA AATGAAGTTT
 5951 TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT
 6001 GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC
 6051 ATAGTTGCCT GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT
 6101 ACCATCTGGC CCCAGTGCTG CAATGATACC GCGAGACCCA CGCTCACC GG
 6151 CTCCAGATTT ATCAGCAATA AACCAGCCAG CCGGAAGGGC CGAGCGCAGA
 6201 AGTGGTCCTG CAACTTTATC CGCCTCCATC CAGTCTATTA ATTGTTGCCG
 6251 GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTTGCGC AACGTTGTTG
 6301 CCATTGCTAC AGGCATCGTG GTGTCACGCT CGTCGTTTGG TATGGCTTCA
 6351 TTCAGCTCCG GTTCCCAACG ATCAAGGCGA GTTACATGAT CCCCCATGTT
 6401 GTGCAAAAAA GCGGTTAGCT CCTTCGGTCC TCCGATCGTT GTCAGAAGTA
 6451 AGTTGGCCGC AGTGTTATCA CTCATGGTTA TGGCAGCACT GCATAATTCT
 6501 CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG GTGAGTACTC
 6551 AACCAAGTCA TTCTGAGAAT AGTGTATGCG GCGACCGAGT TGCTCTTGCC
 6601 CGGCGTCAAT ACGGGATAAT ACCGCGCCAC ATAGCAGAAC TTAAAAGTG
 6651 CTCATCATTG GAAAACGTTT TTCGGGGCGA AAACCTCTCA GGATCTTACC
 6701 GCTGTTGAGA TCCAGTTCGA TGTAACCCAC TCGTGCACCC AACTGATCTT
 6751 CAGCATCTTT TACTTTCACC AGCGTTTCTG GGTGAGCAAA AACAGGAAGG
 6801 CAAAATGCCG CAAAAAAGG AATAAGGGCG ACACGGAAAT GTTGAATACT
 6851 CATACTCTTC CTTTTTCAAT ATTATTGAAG CATTTATCAG GGTATTGTC
 6901 TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA ACAAATAGGG
 6951 GTTCCGCGCA CATTTCCCCG AAAAGTGCCA CCTGACGTCG ACGGATCGGG

Tabelle 11: Nukleotidsequenz des rekombinanten Plasmids pIND-H-SemaL-EA (SEQ ID NO.: 38)

1 AGATCTCGGC CGCATATTAA GTGCATTGTT CTCGATACCG CTAAGTGCAT
 51 TGTCTCTGTT AGCTCGATGG ACAAGTGCAT TGTCTCTTGG CTGAAAGCTC
 101 GATGGACAAG TGCATTGTTT TCTTGCTGAA AGCTCGATGG ACAAGTGCAT
 151 TGTCTCTTGG CTGAAAGCTC AGTACCCGGG AGTACCCTCG ACCGCCGGAG
 201 TATAAATAGA GGCGCTTCGT CTACGGAGCG ACAATTCAAT TCAAACAAGC

251 AAAGTGAACA CGTCGCTAAG CGAAAGCTAA GCAAATAAAC AAGCGCAGCT
 301 GAACAAGCTA AACAATCTGC AGTAAAGTGC AAGTTAAAGT GAATCAATTA
 5 351 AAAGTAACCA GCAACCAAGT AAATCAACTG CAACTACTGA AATCTGCCAA
 401 GAAGTAATTA TTGAATACAA GAAGAGAACT CTGAATACTT TCAACAAGTT
 451 ACCGAGAAAG AAGAACTCAC ACACAGCTAG CGTTTAAACT TAAGCTTGGT
 501 ACCGAGCTCG GATCCACTAG TCCAGTGTGG TGgaattcgg cttgggatga
 10 551 cgcctcctcc gcccgagcgt gccgccccca gcgcaccgcg cgcccgcgtc
 601 cctggccccgc cggctcgggtt ggggcttccg ctgcggctgc ggctgctgct
 651 gctgctctgg gcggccgccc cctccgcccc gggccacctt aggagcggac
 701 cccgcattctt cgccgtctgg aaaggccatg tagggcagga ccgggtggac
 15 751 tttggccaga ctgagccgca cacggtgctt ttccacgagc caggcagctc
 801 ctctgtgtgg gtgggaggac gtggcaaggc ctacctcttt gacttccccg
 851 agggcaagaa cgcatctgtg cgcacggtga atatcggtc cacaaggagg
 20 901 tcctgtctgg ataagcggga ctgcgagaac tacatcactc tcctggagag
 951 gcggagtggg gggctgctgg cctgtggcac caacgcccgg caccagct
 1001 gctggaacct ggtgaatggc actgtggtgc cacttggcga gatgagaggc
 1051 tacgccccct tcagcccgga cgagaactcc ctggttctgt ttgaagggga
 25 1101 cgagggtgtat tccaccatcc ggaagcagga atacaatggg aagatccctc
 1151 ggttccgccc catccggggc gagagtgagc tgtacaccag tgatactgtc
 1201 atgcagaacc cacagttcat caaagccacc atcgtgcacc aagaccaggc
 1251 ttacgatgac aagatctact acttcttccg agaggacaat cctgacaaga
 30 1301 atcctgaggc tcctctcaat gtgtcccgtg tggcccagtt gtgcaggggg
 1351 gaccaggggtg gggaaagttc actgtcagtc tccaagtggg acacttttct
 1401 gaaagccatg ctggtatgca gtgatgctgc caccaacaag aacttcaaca
 1451 ggctgcaaga cgtcttctctg ctccctgacc ccagcggcca gtggaggggac
 35 1501 accaggggtct atgggtgttt ctccaacccc tggaactact cagccgtctg
 1551 tgtgtattcc ctccgtgaca ttgacaaggc cttccgtacc tcctcactca
 1601 agggctacca ctcaagcctt cccaaccgcg ggcctggcaa gtgcctccca
 40 1651 gaccagcagc cgataccac agagaccttc cagggtggctg accgtcaccc
 1701 agagggtggc gagagggtgg agcccatggg gcctctgaag acgccattgt
 1751 tccactctaa ataccactac cagaaagtgg ccgttcaccg catgcaagcc
 1801 agccacgggg agaccttca tgtgctttac ctaactacag acaggggcac
 45 1851 tatccacaag gtggtggaac cgggggagca ggagcacagc ttgccttca
 1901 acatcatgga gatccagccc ttccgcccgc cggtgccat ccagaccatg
 1951 tcgctggatg ctgagcggag gaagctgtat gtgagctccc agtgggaggc
 2001 gagccaggtg cccctggacc tgtgtgaggt ctatggcggg ggctgccacg
 50 2051 gttgcctcat gtcccagac ccctactgcg gctgggacca gggccgctgc
 2101 atctccatct acagctccga acggtcagtg ctgcaatcca ttaatccagc
 2151 cgagccacac aaggagtgtc ccaaccccaa accagacaag gcccactgc

2201 agaaggtttc cctggcccca aactctcgct actacctgag ctgccccatg
 2251 gaatcccgcc acgccaccta ctcatggcg cacaaggaga acgtggagca
 5 2301 gagctgcgaa cctggtcacc agagcccca ctgcatactg ttcacgcaga
 2351 acctcacggc gcagcagtag ggccactact tctgcgaggc ccaggagggc
 2401 tcctacttcc gcgaggctca gcaactggcag ctgctgcccg aggacggcat
 2451 catggccgag cacctgctgg gtcattgcctg tgccctggct gcctccctct
 10 2501 ggctgggggt gctgcccaca ctactcttg gcttgctggt ccacgtgaag
 2551 cttGGGCCCC AACAAAACT CATCTCAGAA GAGGATCTGA ATAGCGCCGT
 2601 CGACCATCAT CATCATCATC ATTGAGTTA TCCAGCACAG TGGCGGCCGC
 2651 TCGAGTCTAG AGGGCCCCGT TAAACCCGT GATCAGCCTC GACTGTGCCT
 15 2701 TCTAGTTGCC AGCCATCTGT TGTTTGGCCC TCCCCCGTGC CTTCTTGAC
 2751 CCTGGAAGGT GCCACTCCCA CTGTCCTTC CTAATAAAAT GAGGAAATTG
 2801 CATCGCATTG TCTGAGTAGG TGTCATTCTA TTCTGGGGGG TGGGGTGGGG
 2851 CAGGACAGCA AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA
 20 2901 TGCGGTGGGC TCTATGGCTT CTGAGGCGGA AAGAACCAGC TGGGGCTCTA
 2951 GGGGGTATCC CCACGCGCCC TGTAGCGGCG CATTAAAGCG GCGGGGTGTG
 3001 GTGGTTACGC GCAGCGTGAC CGCTACACTT GCCAGCGCCC TAGCGCCCCG
 3051 TCCTTTCGCT TTCTTCCCTT CCTTCTCGC CACGTTCCGC GGCTTTCCTC
 25 3101 GTCAAGCTCT AAATCGGGGC ATCCCTTTAG GGTTCGATT TAGTGCTTTA
 3151 CGGCACCTCG ACCCCAAAAA ACTTGATTAG GGTGATGGT CACGTAGTGG
 3201 GCCATCGCCC TGATAGACGG TTTTTCGCCC TTTGACGTTG GAGTCCACGT
 30 3251 TCTTTAATAG TGGACTCTTG TTCCAACTG GAACAACACT CAACCCTATC
 3301 TCGGTCTATT CTTTGTGATT ATAAGGGATT TTGGGGATT CGGCCTATTG
 3351 GTTAAAAAAT GAGCTGATTT AACAAAAAT TAACGCGAAT TAATTCTGTG
 3401 GAATGTGTGT CAGTTAGGGT GTGGAAAGTC CCCAGGCTCC CCAGGCAGGC
 35 3451 AGAAGTATGC AAAGCATGCA TCTCAATTAG TCAGCAACCA GGTGTGGAAA
 3501 GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG CATCTCAATT
 3551 AGTCAGCAAC CATAGTCCCG CCCCTAACTC CGCCCATCCC GCCCTAACT
 3601 CCGCCAGTT CCGCCCATTC TCCGCCCAT GGCTGACTAA TTTTTTTTAT
 40 3651 TTATGCAGAG GCCGAGGCCG CCTCTGCCTC TGAGCTATTC CAGAAGTAGT
 3701 GAGGAGGCTT TTTTGGAGGC CTAGGCTTTT GCAAAAAGCT CCCGGGAGCT
 3751 TGTATATCCA TTTTCGGATC TGATCAAGAG ACAGGATGAG GATCGTTTCG
 3801 CATGATTGAA CAAGATGGAT TGCACGCAGG TTCTCCGGCC GCTTGGGTGG
 45 3851 AGAGGCTATT CGGCTATGAC TGGGCACAAC AGACAATCGG CTGCTCTGAT
 3901 GCCGCCGTGT TCCGGCTGTC AGCGCAGGGG CGCCCGGTTT TTTTGTCAA
 3951 GACCGACCTG TCCGGTGCCC TGAATGAACT GCAGGACGAG GCAGCGCGGC
 50 4001 TATCGTGGCT GGCCACGACG GCGGTTCTTT GCGCAGCTGT GCTCGACGTT
 4051 GTCAGTGAAG CGGGAAGGGA CTGGCTGCTA TTGGGCGAAG TGCCGGGGCA
 4101 GGATCTCCTG TCATCTCACC TTGCTCCTGC CGAGAAAGTA TCCATCATGG

4151 CTGATGCAAT GCGGCGGCTG CATACGCTTG ATCCGGCTAC CTGCCCATTG
 4201 GACCACCAAG CGAAACATCG CATCGAGCGA GCACGTACTC GGATGGAAGC
 5 4251 CGGTCTTGTC GATCAGGATG ATCTGGACGA AGAGCATCAG GGGCTCGCGC
 4301 CAGCCGAAC TTTCCGCCAGG CTCAAGGCGC GCATGCCCCG CGGCGAGGAT
 4351 CTCGTCGTGA CCCATGGCGA TGCCTGCTTG CCGAATATCA TGGTGGAAAA
 4401 TGGCCGCTTT TCTGGATTCA TCGACTGTGG CCGGCTGGGT GTGGCGGACC
 10 4451 GCTATCAGGA CATAGCGTTG GCTACCCGTG ATATTGCTGA AGAGCTTGGC
 4501 GCGGAATGGG CTGACCGCTT CCTCGTGCTT TACGGTATCG CCGCTCCCGA
 4551 TTCGAGCGC ATCGCCTTCT ATCGCCTTCT TGACGAGTTC TTCTGAGCGG
 4601 GACTCTGGGG TTCGAAATGA CCGACCAAGC GACGCCCCAAC CTGCCATCAC
 15 4651 GAGATTTTCA TTCCACCGCC GCCTTCTATG AAAGGTTGGG CTTCCGGAATC
 4701 GTTTTCCGGG ACGCCGGCTG GATGATCCTC CAGCGCGGGG ATCTCATGCT
 4751 GGAGTTCTTC GCCCACCCCA ACTTGTTTAT TGCAGCTTAT AATGGTTACA
 4801 AATAAAGCAA TAGCATCACA AATTTACAA ATAAAGCATT TTTTTCAC TG
 20 4851 CATTCTAGTT GTGGTTTGTC CAAACTCATC AATGTATCTT ATCATGTCTG
 4901 TATACCGTCG ACCTCTAGCT AGAGCTTGGC GTAATCATGG TCATAGCTGT
 4951 TTCCTGTGTG AAATTGTTAT CCGCTCACAA TTCCACACAA CATACGAGCC
 5001 GGAAGCATAA AGTGTAAGC CTGGGGTGCC TAATGAGTGA GCTAACTCAC
 25 5051 ATTAATTGCG TTGCGCTCAC TGCCCGCTTT CCAGTCGGGA AACCTGTCTG
 5101 GCCAGCTGCA TTAATGAATC GGCCAACGCG CGGGGAGAGG CGGTTTGC GT
 5151 ATTGGGCGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCTG
 30 5201 TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT
 5251 CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
 5301 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG
 5351 GCTCCGCCCC CCTGACGAGC ATCACA AAAA TCGACGCTCA AGTCAGAGGT
 35 5401 GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC
 5451 TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC
 5501 CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC TCACGCTGTA
 5551 GGTATCTCAG TTCGGTGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC
 40 5601 GAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT
 5651 TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
 5701 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG
 5751 AAGTGGTGGC CTAAC TACGG CTACACTAGA AGGACAGTAT TTGGTATCTG
 45 5801 CGCTCTGCTG AAGCCAGTTA CCTTCGGA AAAGAGTTGGT AGCTCTTGAT
 5851 CCGGCAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG
 5901 CAGATTACGC GCAGAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC
 5951 TACGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTAA GGGATTTTGG
 50 6001 TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT AAATTAAAAA
 6051 TGAAGTTTAA AATCAATCTA AAGTATATAT GAGTAACTT GGTCTGACAG

6101 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTT
 6151 GTTCATCCAT AGTTGCCTGA CTCCCCGTCG TG TAGATAAC TACGATACGG
 5 6201 GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG
 6251 CTCACCGGCT CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG
 6301 AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAAT
 6351 TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA
 10 6401 CGTTGTTGCC ATTGCTACAG GCATCGTGGT GTCACGCTCG TCGTTTGGA
 6451 TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC
 6501 CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCCTC CGATCGTTGT
 6551 CAGAAGTAAG TTGGCCGAG TGTTATCACT CATGGTTATG GCAGCACTGC
 15 6601 ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT
 6651 GAGTACTCAA CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG
 6701 CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT AGCAGAACTT
 6751 TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG
 20 6801 ATCTTACCGC TGTTGAGATC CAGTTCGATG TAACCCACTC GTGCACCCAA
 6851 CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA
 6901 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGCGGAC ACGGAAATGT
 6951 TGAATACTCA TACTCTTCCT TTTTCAATAT TATTGAAGCA TTTATCAGGG
 25 7001 TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC
 7051 AAATAGGGGT TCCGCGCACA TTTCCCGAA AAGTGCCACC TGACGTCGAC
 7101 GGATCGGG

30

Tabelle 12: Sequenz des rekombinanten Plasmids pQE30-H-SemaL-BH (SEQ ID NO.: 39)

35 1 CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT
 51 AATAGATTCA ATGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG
 101 AGGAGAAATT AACTATGAGA GGATCGCATC ACCATCACCA TCACGGatcc
 151 ctggttctgt ttgaagggga cgaggtgtat tccaccatcc ggaagcagga
 40 201 atacaatggg aagatccctc ggttccgccc catccggggc gagagtgagc
 251 tgtacaccag tgatactgtc atgcagaacc cacagttcat caaagccacc
 301 atcgtgcacc aagaccaggc ttacgatgac aagatctact acttcttccg
 351 agaggacaat cctgacaaga atcctgagge tcctctcaat gtgtcccgtg
 45 401 tggcccagtt gtgcaggggg gaccaggggtg gggaaagtgc actgtcagtc
 451 tccaagtgga acacttttct gaaagccatg ctggtatgca gtgatgctgc
 501 caccaacaag aacttcaaca ggctgcaaga cgtcttcctg ctcccctgacc
 551 ccagcggcca gtggagggac accaggggtc atggtgtttt ctccaacccc
 50 601 tggaactact cagccgtctg tgtgtattcc ctcggtgaca ttgacaaggt
 651 cttccgtacc tcctcactca agggctacca ctcaagcctt cccaacccgc

55

701 ggccctggcaa gtgcctccca gaccagcagc cgataccac agaAAGCTTA
 751 ATTAGCTGAG CTTGGACTCC TGTGATAGA TCCAGTAATG ACCTCAGAAC
 801 TCCATCTGGA TTTGTTCCAGA ACGCTCGGTT GCCGCCGGGC GTTTTTTATT
 851 GGTGAGAATC CAAGCTAGCT TGGCGAGATT TTCAGGAGCT AAGGAAGCTA
 901 AAATGGAGAA AAAAATCACT GGATATACCA CCGTTGATAT ATCCCAATGG
 951 CATCGTAAAG AACATTTTGA GGCATTTTCAG TCAGTTGCTC AATGTACCTA
 1001 TAACCAGACC GTTCAGCTGG ATATTACGGC CTTTTTAAAG ACCGTAAAGA
 1051 AAAATAAGCA CAAGTTTTAT CCGGCCTTTA TTCACATTCT TGCCCCGCTG
 1101 ATGAATGCTC ATCCGGAATT TCGTATGGCA ATGAAAGACG GTGAGCTGGT
 1151 GATATGGGAT AGTGTTCACC CTTGTACAC CGTTTTCCAT GAGCAAACCTG
 1201 AAACGTTTTC ATCGCTCTGG AGTGAATACC ACGACGATTT CCGGCAGTTT
 1251 CTACACATAT ATTCGCAAGA TGTGGCGTGT TACGGTGAAA ACCTGGCCTA
 1301 TTTCCCTAAA GGGTTTATTG AGAATATGTT TTTCGTCTCA GCCAATCCCT
 1351 GGGTGAGTTT CACCAGTTTT GATTTAAACG TGGCCAATAT GGACAACTTC
 1401 TTCGCCCCCG TTTTCACCAT GGGCAAATAT TATACGCAAG GCGACAAGGT
 1451 GCTGATGCCG CTGGCGATTG AGGTTTCATCA TGCCGTCTGT GATGGCTTCC
 1501 ATGTCGGCAG AATGCTTAAT GAATTACAAC AGTACTGCGA TGAGTGGCAG
 1551 GGCGGGGCGT AATTTTTTTA AGGCAGTTAT TGGTGCCCTT AAACGCCTGG
 1601 GGTAATGACT CTCTAGCTTG AGGCATCAA TAAAACGAAA GGCTCAGTCG
 1651 AAAGACTGGG CCTTTCGTTT TATCTGTTGT TTGTCGGTGA ACGCTCTCCT
 1701 GAGTAGGACA AATCCGCCGC TCTAGAGCTG CCTCGCGCGT TTCGGTGATG
 1751 ACGGTGAAAA CCTCTGACAC ATGCAGCTCC CGGAGACGGT CACAGCTTGT
 1801 CTGTAAGCGG ATGCCGGGAG CAGACAAGCC CGTCAGGGCG CGTCAGCGGG
 1851 TGTTGGCGGG TGTCGGGGCG CAGCCATGAC CCAGTCACGT AGCGATAGCG
 1901 GAGTGTATAC TGGCTTAACT ATGCGGCATC AGAGCAGATT GTACTGAGAG
 1951 TGCACCATAT GCGGTGTGAA ATACCGCACA GATGCGTAAG GAGAAAATAC
 2001 CGCATCAGGC GCTCTTCCGC TTCCTCGCTC ACTGACTCGC TGCGCTCGGT
 2051 CTGTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT
 2101 TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC
 2151 CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA
 2201 TAGGCTCCGC CCCCTGACG AGCATCACA AAATCGACGC TCAAGTCAGA
 2251 GGTGGCGAAA CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA
 2301 AGCTCCCTCG TGCGCTCTCC GTTCCGACC CTGCCGCTTA CCGGATACCT
 2351 GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCACGCT
 2401 GTAGGTATCT CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG
 2451 CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATATCG
 2501 TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA
 2551 CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC
 2601 TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG TATTTGGTAT

2651 CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT
 2701 GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG
 2751 CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT
 2801 TTCTACGGGG TCTGACGCTC AGTGAACGA AAACACACGT TAAGGGATTT
 2851 TGGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA
 2901 AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA
 2951 CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT
 3001 TTCGTTTCATC CATAGCTGCC TGAATCCCCG TCGTGTAGAT AACTACGATA
 3051 CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC
 3101 ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG
 3151 CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT
 3201 AATTGTTGCC GGGAAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG
 3251 CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG
 3301 GTATGGCTTC ATTACAGCTCC GGTTCCTAAC GATCAAGGCG AGTTACATGA
 3351 TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT
 3401 TGTCAGAAGT AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC
 3451 TGCATAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT
 3501 GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCGACCGAG
 3551 TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA
 3601 CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA
 3651 AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC
 3701 CAACTGATCT TCAGCATCTT TTACTTTTAC CAGCGTTTCT GGGTGAGCAA
 3751 AAACAGGAAG GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA
 3801 TGTTGAATAC TCATACTCTT CCTTTTTTCAA TATTATTGAA GCATTTATCA
 3851 GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA
 3901 AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC
 3951 TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC
 4001 GAGGCCCTTT CGTCTTCAC

Tabelle 13: Sequenz des rekombinanten Plasmids pQE31-H-SemaL-SH (SEQ ID NO.: 40)

1 CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT
 51 AATAGATTCA ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG
 101 AGGAGAAATT AACTATGAGA GGATCGCATC ACCATCACCA TCACACGGAT
 151 CCGCATGCga gctcccagtg ggaggtgagc caggtgcccc tggacctgtg
 201 tgaggtctat ggcgggggct gccacggttg cctcatgtcc cgagaccctt
 251 actgcggttg ggaccagggc cgctgcatct ccatctacag ctccgaacgg
 301 tcagtgtctgc aatccattaa tccagccgag ccacacaagg agtgtcccaa

351 ccccaaacca gacaaggccc cactgcagaa ggtttccctg gccccaaact
 401 ctcgctacta cctgagctgc cccatggaat cccgccacgc cacctactca
 5 451 tggcgccaca aggagaacgt ggagcagagc tgcgaacctg gtcaccagag
 501 ccccaactgc atcctgttca tcgagaacct cacggcgag cagtacggcc
 551 actacttctg cgaggcccag gagggctcct acttccgcga ggctcagcac
 601 tggcagctgc tgcccagagga cggcatcatg gccgagcacc tgctgggtca
 10 651 tgccgtgtgcc ctggctgcct ccctctggct gggggtgctg cccacactca
 701 ctcttggtt gctgggtccac gtgaagctta ATTAGCTGAG CTTGGACTCC
 751 TGTTGATAGA TCCAGTAATG ACCTCAGAAC TCCATCTGGA TTTGTTTCTG
 801 ACGCTCGGTT GCCGCCGGGC GTTTTTTATT GGTGAGAATC CAAGCTAGCT
 15 851 TGGCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA AAAAATCACT
 901 GGATATACCA CCGTTGATAT ATCCCAATGG CATCGTAAAG AACATTTTGA
 951 GGCATTTTCTG TCAGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG
 1001 ATATTACGGC CTTTTTAAAG ACCGTAAAGA AAAATAAGCA CAAGTTTTAT
 20 1051 CCGGCCTTTA TTCACATTCT TGCCCGCCTG ATGAATGCTC ATCCGGAATT
 1101 TCGTATGGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT AGTGTTCACC
 1151 CTTGTTACAC CGTTTTCCAT GAGCAAACCTG AAACGTTTTT ATCGCTCTGG
 1201 AGTGAATACC ACGACGATTT CCGGCAGTTT CTACACATAT ATTCGCAAGA
 25 1251 TGTGGCGTGT TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG
 1301 AGAATATGTT TTTCTGCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT
 1351 GATTTAAACG TGGCCAATAT GGACAACTTC TTCGCCCCCG TTTTCACCAT
 1401 GGGCAAATAT TATACGCAAG GCGACAAGGT GCTGATGCCG CTGGCGATTC
 30 1451 AGGTTTCATCA TGCCGTCTGT GATGGCTTCC ATGTCGGCAG AATGCTTAAT
 1501 GAATTACAAC AGTACTGCGA TGAGTGGCAG GGCGGGGCGT AATTTTTTTA
 1551 AGGCAGTTAT TGGTGCCCTT AAACGCCTGG GGTAATGACT CTCTAGCTTG
 35 1601 AGGCATCAAA TAAAACGAAA GGCTCAGTCG AAAGACTGGG CCTTTCTGTTT
 1651 TATCTGTTGT TTGTGCGTGA ACGCTCTCCT GAGTAGGACA AATCCGCCGC
 1701 TCTAGAGCTG CCTCGCGCGT TTCGGTGATG ACGGTGAAAA CCTCTGACAC
 1751 ATGCAGCTCC CGGAGACGGT CACAGCTTGT CTGTAAGCGG ATGCCGGGAG
 40 1801 CAGACAAGCC CGTCAGGGCG CGTCAGCGGG TGTGCGGGG TGTCGGGGCG
 1851 CAGCCATGAC CCAGTCACGT AGCGATAGCG GAGTGTATAC TGGCTTAACT
 1901 ATGCGGCATC AGAGCAGATT GTAGTGAGAG TGCACCATAT GCGGTGTGAA
 1951 ATACCGCACA GATGCGTAAG GAGAAAATAC CGCATCAGGC GCTCTTCCGC
 45 2001 TTCCTCGCTC ACTGACTCGC TGCGCTCGGT CTGTCGGCTG CGGCGAGCGG
 2051 TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT
 2101 AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG CCAGGAACCG
 2151 TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCTGACG
 50 2201 AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA
 2251 CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC

2301 TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG
 2351 GAAGCGTGGC GCTTTCTCAA TGCTCACGCT GTAGGTATCT CAGTTCGGTG
 2401 TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC
 2451 CGACCGCTGC GCCTTATCCG GTAACATATCG TCTTGAGTCC AACCCGGTAA
 2501 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA
 2551 GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA
 2601 CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG
 2651 TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC
 2701 GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA CGCGCAGAAA
 2751 AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC
 2801 AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA
 2851 AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT
 2901 CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA
 2951 GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTTCATC CATAGCTGCC
 3001 TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG
 3051 CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT
 3101 TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT
 3151 GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG
 3201 AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA
 3251 CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTACAGCTCC
 3301 GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCATGT TGTGCAAAAA
 3351 AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAAGT AAGTTGGCCG
 3401 CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC
 3451 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
 3501 ATTCTGAGAA TAGTGTATGC GCGGACCGAG TTGCTCTTGC CCGGCGTCAA
 3551 TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT
 3601 GGAAAACGTT CTTGCGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG
 3651 ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT
 3701 TTACTTTTAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC
 3751 GCAAAAAAGG GAATAAGGGC GACACGAAA TGTTGAATAC TCATACTCTT
 3801 CCTTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG
 3851 GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCGCGC
 3901 ACATTTCCCC GAAAAGTGCC ACCTGACGTC TAAGAAACCA TTATTATCAT
 3951 GACATTAACC TATAAAAATA GCGGTATCAC GAGGCCCTTT CGTCTTCAC

Tabelle 14: (Partielle) Nukleotidsequenz des humanen Semaphorin L Gens.

(8888 Nukleotide) (SEQ ID NO.: 41);

GAGCCGCACACGGTGCTTTTCCACGAGCCAGGCAGCTCCTCTGTGTGGGTGGGAGGACGT
GGCAAGGTCTACCTCTTTGACTTCCCCGAGGGCAAGAACGCATCTGTGCGCACGGTGAGC
5 CTCTCTCTTCCCCAACACCCCCCTACCCTCTTATCTCCCCTCTGGCCCTGCCAAGGGT
CCTCAGGGAATCCGAGGGAGCTGGCTTCTCTTCTAAACTGCCCCACCTCCGTATCCTA
TAAATGGCTCCTGGGGGAGGCTCCCTAAAGGTAGTCCAGATTGGAGTGGGGAGCTGGGGC
GGTGTGGAGAAAAACAGGAGCTAATGGGCCTGGCCAGCTGGGCAGCGCTGCTGCGGAAAG
10 CCCAGGCTGGAAGCTGGGCCCCAGAGCCCATGCCTGGTCTTCTGAACCCTCTGGGCCTCA
GCTCTGGATATGAGACCCTGTTTGACCTCAGGTAGATCACTCACCCTCTCAGAGCCCCAG
TTGCTCATCTGTGAGATGAGAATAATGGTTGCTTCTTTGGGGCTTATCCTGAGGCTGTG
TGGAAAGCATTTTCAGGGGTACCTCACCCCTGGCAGATTGAACTAATGCTTCTCCCCTTCC
15 CCAGGTGAATATCGGCTCCACAAAGGGGTCTGTCTGGATAAGCGGGTGAGCGGGGGAGG
GATCTGGAGGGGTCTGAGCCACTTGGTAAAGGGAGAGGAGACCCTGAGGGTCTAAGGAAG
GAAGCATGGCCCTGCCCCACGAGTCCAGACTGATGGGGAGACGTGGTCTCTGTGCTTA
GGGGATGGCGTCAGCTGCACACACTCTGGGCTGTCCCGGGAGGCTGTCACCTATGCTAAG
20 CCCTTCTGACACCTTCTTCCCTGATCCTGGGGGTCTAGTGCTAGGCTTGCCAGGGCCTT
CCAGCAACCAATTTCTCTCCTCCCTTCTCTCTTCCCCGGGCAGGACTGCGAGAATAACAT
CACTCTCCTGGAGAGGCGGAGTGAGGGGCTGCTGGCCTGTGGCACCAACGCCCGGCACCC
CAGCTGCTGGAACCTGGTGAGAAGGCTGCTCCCCATGTGCCTGATCAGCTCACCTTCTAC
25 TGCGTGGGCTTCTGCCCTCATGGTGGGAAGGAGATGGCGAGACTCCAATGCTGGCCTTG
CCCTGGGAGGATGGGGCTCCTGGCCGAGAACTGGCCGTATGGGAGGCAGTGGCTGTGG
GATTATGTGGCCATCCAACCCTCTGGATCTCCACAGGTGAATGGCACTGTGGTGCCACT
TGCGGAGATGAGAGGCTACGCCCCCTTCAGCCCGGACGAGAACTCCCTGGTTCTGTTTGA
30 AGGTTGGGGCATGCTTCGGAACCTGGGCTGGGAGCAGGATGGTCAGCTCTTTGTCCAGTGT
CCGGAGGAGGGACTTCCAGGAGCTGCCTGCCCTTACTCATTCTCCCTCCCACTGACCCC
AGGGGACGAGGTGTATTCCACCATCCGGAAGCAGGAATACAATGGGAAGATCCCTCGGTT
35 CCGCCGCATCCGGGGCAGAGTGAGCTGTACACAGTGATACTGTGATGCAGAGTGAGTC
AGGCTCCGGCTGGGCTGAGGGTGGGCAAGGGGGTGTGAGCACTTAAGGTGGCAGATGGGA
TCCTGATGTTTCTGGGAGGGCTCCCTGAGGGCCGCTGGGGCCATGCAGGAAAGCAGGACC
TTGGTATAGGCCTGAGAAGTTAGGGTTGGCTGGGAGCAGAGGAACAGACAAGGTATAGCA
40 GTGGGATGGGCCCAGCCCTCTTCAGGAACACAAACAGAGGGAGCCCCAGACCCAGTGCAG
GGTCCCCAGGAGCCAAAGTTTATCCTCTGCTGAGTTCACGTGGAGGCAGCCCCCAACTC
CCTCCTCATCAGGGCTCTGCCAATTGAGCAGAAGTGACATAGGGGCCCCCAGGGACCTTC
CCCCACTCCCCAGGCATGAAGTCATTGCTCCTGGGCCGATGACATCTTTGTAGGAAGAGG
45 GCAAAACAGGTGTGGGGTGGAGGTGCAGGGTCTAGGGCCCCCTCGGGGAGTTGGACCTGAT
GTTATGAGTCCTATTCCAGATCTGATTTGCCATGGTTTGTGCAGACCCGAAGGAGGGAGG
AGAGTGTGCAGGGTTGGAATGGTCTCCCGGGCAAGCTTCCCAGCCTTACGCCCATTCGCT
50 TCTGTGCCCTGGCAGACCCACAGTTCATCAAAGCCACCATCGTGCACCAAGACCAGGCTT
ACGATGACAAGATCTACTACTTCTTCCGAGAGGACAACTCCTGACAAGAATCCTGAGGCTC
CTCTCAATGTGTCCCGTGTGGCCAGTTGTGCAGGGTGAACACGGGCGTGAGGGCTGCTG

5
10
15
20
25
30
35
40
45
50

GGAGGGGCCTGAACCACAGGGATACTGGCCTCTGCCAGGCAGAATGAGGGAGTCAGGCCC
TGCGCCTGTCTTTGGGATTGTGCAGGTGAGAAGAAACATTTGAGGAGTTGATGGGGCACA
AATTAGGTATGGGGAAGGAGTTCCAGGGGGCAGAACCTTTGCCATCTCACAGAGGACAGG
GGCAGCTTCTCTTCTTCCCTGGAGTAGGCCCTGCTGGGGGAAGCTGGGTGGAATGCCGTG
GGAGATGCTCCTGCTTCTTGGAAAGCCACAGGACACGGAGGAGCCAGTCTGAGTTGGGT
TTGTGCGCAGCTTCCCATGCCAGCTGCCTTCCTTGAGACTGGAAAGGGCCTCTAGCACCCC
TGGGGCCATTCAATTCAGGCCCAGGCGCCCAACCTCAGTTGTTACATTCCCCATGTGAT
CTCCTGTTGCTGCTTCACCTTGGGACTGTCTCGGCTTTGGTGACCTTGAGGAACTGGA
ACCCACAGCACCATTGTTTGGCTCCTGGAAGCCTTGGGGAGAGGAATTTCCACAGGGCAG
GGCCTGGGTCTGATTCCCTGCCTCTTTACTCCCTATTTCATCCCGGCTACACCCTTGGGC
CCCCATCCTTGCTTGGCTCCAGTACTGGCTGGCACAGCTGTTGTGGTCATCCAGGGATGG
CAGGGCACTGGGGAACAGAAGAGAGAGGTACACAGTGCGGAACCTGGGAGCAGGAGCTAG
GACAAGGAAGGCTGGACTTGGGCCATGGATTCCCTTCCTGCAGACTTGGGAAGTGAGCAC
ACTTGAGTGATTAGAGAAGGTGTCTTCGTTCTAAGGGCAGTGGAGGAGGCACCATTTTGG
AGCCTGCATCATTTCGTATTTGGGCTAGATTGAAAAATAGAGCTTTCTAAGTCTCTGCAG
AGAATGGGAGGCTCTCACAACTGGGAGAAGTATTGGCTCTTTTCTGAGAATTTTGCCAA
GGGTATGCTGTACTGGGGCTGGTTTGAAGGAGTATAGGGCATTATGTCTGTGAAGGCA
GTGGCTGGGGTGGGGCCTTATCAGGCCCAAGGAGCATCTGGCCACATCTCAGAGTCCACA
GATGAGGATCACGGATGTGTAGAGGAAACATCCTAGGCAGGCAATCATCTGACTGCTTTT
TTGGGGCAGGTGATGCCCTGGGAAATTGGGAGGGAGGGAGAGGGAGGTAGGCTATTCT
AGAACTGGGAGAGCAGGTGAGGTAGGATTGGGAGGACCAGGGGTCAGGGTCCCCATTGG
TCCCTAATTGAGAACGGAGAGAGCATTGGTCTAGGAGGCAGGCAGCTCGGTTATAAGACC
TTGGGAACCTCTTGATTTAGAATCCAAGATCCTTTTTTAGATCTAGGATTTTATAAAATTAA
GATATCCCCTAAGATCAAATGCAACGTGGAGTCTCTGAATTGGATCCTAGAACAGAAGAAG
GACATTTGTGAAAAAACTAGTGAAATCCAAATAAAGTCTGTAGTTTTGTTAATAGTAATG
CACCAATGTCAGTTGCCTAGTTGTGACAAATATAACGTGGTTATGTAAGATGGTAACATT
AGGGGGAACCTGGAGAAGGGTAGATTGGAGCTCTCTGTACTATCTTTGCAACTTTTCTGGG
AATCTAAAATTACTCCAAAATAAAAAAAAATGTATTTAAAGTAAATATATTCCCTAAGA
GTCCAGGAGGCAGGGGAGTTGTAGAAGCAGCTGAGTGGTTGGGTCTGACAGATTTGGTT
CCAACTCGGTCTCTGCTGCTCACCAGCTGTGTGACCTTGAGCAAGTGGCTTAGCCTTTCT
GAGCCTGATTTCTTTATCTGTGGAGTGGGGAAGATGACAGCCACCTCGCAGGGCTGTGGA
GGGTAAACGAGGTGATGCATGGACAGCAGCCGCACTGACCTTGCTGGTGTGGGGCTCCT
GCTTCTGTTCTTCCCGTGCAGCCTTGGGAATGTTGGAGGCCGTATCCAGGGACCCCTGGG
CCTCCTGGGATGGCCTCTCTGGATCAGCCTTGAAGGTTCCAGGCTGCCCTTAGGCTCCC
ACATTCTTCCCCAGTCACGCTCTCCTCGCCCTGCCCACACCAGTCTGTGACCTTGCCT
GAGTTGTGACTTCCCACCCCTCCCCGGCCTAGAGGAAAGCTGCCTGGCCCCCTCAGTGGGA
CTCCCGCCCACTGACCTCTGTCCACCATAACAGACAGGGGCACTATCCACAAGGTGGT
GGAACCGGGGAGCAGGAGCACAGCTTCGCCTTCAACATCATGGAGATCCAGCCCTTCCG
CCGCGCGGCTGCCATCCAGACCATGTCGCTGGATGCTGAGCGGGTGAGCCTTCCCCACT

GCGTCCCATGGGCTATGCAGTGAAGTGCAGCTGAGGACAGGGCTCCTTTGCATGTGATTTG
 TGTGTTCTTTTAAGAGCTTCTAGGCCTTAGGGCCTGGACATTTAGGACTGAGTGTGGGT
 5 GGGGCCCCGGGCTGACCAATCCTGCTGTCTTCCAGAGGAAGCTGTATGTGAGCTCCCA
 GTGGGAGGTGAGCCAGGTGCCCTGGACCTGTGTGAGGTCTATGGCGGGGGCTGCCACGG
 TTGCCTCATGTCCCGAGACCCCTACTGCGGCTGGGACCAGGGCCGCTGCATCTCCATCTA
 CAGCTCCGAACGGTACGTTGGCCGGGATCCCTCCGTCCCTGGGACAAGGTGGGCATGGGA
 10 CAGGGGGAGGTGTTGTGCGGCTGGAAGAGGTGGCGGTACTGGGCCTTTCTTGTGGGACCT
 CCTCTCTACTGGAAGTGCAGTGGGGTAAGGATATGAGGGTCAGGTCTGCAGCCTTGTAT
 CTGTGATCCTCTTTTCGTCTTCCCACTCCAGGTCAGTGCTGCAATCCATTAATCCAGCC
 GAGCCACACAAGGAGTGTCCCAACCCCAAACCAGGTACCTGATCTGGCCCTGCTGGCGGC
 15 TGTGGCCCAATGAGTGGGGTACTGCCCTGCCCTGATTGTCTGGTCTGAGGGAAACATGG
 CCTTGTCTGTGGGCCCCAGGTACATGGGGCAGGATACAGTCTGCAGAGGGAGCCCTCT
 TGGTGGGATGAGCGAGACGGGAGAAAAAAGGAGGACGCTGAGGGCTGGGTTCACCGTT
 CATTGAGAAGCCTTGTCTGGGATCCCAGTCGGTGGGAGGACACATCCTCCCTGGGAG
 20 CTCTTTGTCCCTCCTCACGGCTGCTTCCCCACTGCCTCCCCAGACAAGGCCCCACTGCAG
 AAGGTTTCCCTGGCCCCAACTCTCGCTACTACCTGAGCTGCCCCATGGAATCCCCGCCAC
 GCCACCTACTCATGGCGCCACAAGGAGAACGTGGAGCAGAGCTGCGAACCTGGTCACCAG
 AGCCCCAACTGCATCCTGTTTCATCGAGAACCTCACGGCGCAGCAGTACGGCCACTACTTC
 25 TGCGAGGCCCAGGAGGGCTCCTACTTCCGCGAGGCTCAGCACTGGCAGCTGCTGCCCGAG
 GACGGCATCATGGCCGAGCACCTGCTGGGTTCATGCCTGTGCCCTGGCCGCTCCCTCTGG
 CTGGGGGTGCTGCCACACTCACTCTTGGCTTGCTGGTCCACTAGGGCCTCCCGAGGCTG
 GGCATGCCTCAGGCTTCTGCAGCCCAGGGCACTAGAACGTCTCACACTCAGAGCCGGCTG
 30 GCGCGGAGCTCCTTGCTGCCACTTCTTCCAGGGGACAGAATAACCCAGTGGAGGATGC
 CAGGCCTGGAGAGCTCCAGCCGAGGCGGCTGCTGGGCCCCAGGTGGCGCACGGATGGTG
 AGGGGCTGAGAATGAGGGCACCGACTGTGAAGCTGGGGCATCGATGACCCAAGACTTTAT
 CTTCTGGAATAATTTTTTCACTCCTCAAACCTTGAATAATGCAGCGATGCTCCAGCC
 35 CAAGAGCCCATGGGTCGGGGAGTGGGTTTGGATAGGAGAGCTGGGACTCCATCTCGACCC
 TGGGGCTGAGGCCTGAGTCCTTCTGGACTCTTGGTACCCACATTGCCTCCTTCCCTCCC
 TCTCTCATGGCTGGGTGGCTGGTGTTCCTGAAGACCCAGGGCTACCCTCTGTCCAGCCCT
 40 GTCCTCTGCAGCTCCCTCTCTGGTCTGGGTCCACAGGACAGCCGCTTGCATGTTTAT
 TGAAGGATGTTTGTCTTCCGGACGGAAGGACGGAAAAAGCTCTGAAAAAAAAAAAAAAAAA
 AAAAAAA

Tabelle 15: Nukleotidsequenz von pMelBacA-HSEMAL (6622bp) (SEQ ID NO: 42)

1 GATATCATGG AGATAATTAA AATGATAACC ATCTCGCAA TAAATAAGTA
 □

50 □ 51 TTTTACTGTT TTCGTAACAG TTTTGTAATA AAAAAACCTA TAAATATGAA

101 ATTCTTAGTC AACGTTGCCC TTGTTTTTAT GGTCTGTATAC ATTTCTTACA
 151 TCTATGCGGA TCGATGG
 5 gga tccgcccagg gccacctaag gagcggaccc
 201 cgcattcttcg ccgtctggaa aggccatgta gggcaggacc ggggtggactt
 251 tggccagact gagccgcaca cgggtgctttt ccacgagcca ggcagctcct
 10 301 ctgtgtgggt gggaggacgt ggcaaggctt acctctttga cttccccgag
 351 ggcaagaacg catctgtgcg cacggtgaat atcgggtcca caaaggggtc
 401 ctgtctggat aagcgggact gcgagaacta catcactctc ctggagaggc
 15 451 ggagtgaggg gctgctggcc tgtggcacca acgcccggca cccagctgc
 501 tggaacctgg tgaatggcac tgtggtgcca cttggcgaga tgagaggcta
 551 tgcccccttc agcccggacg agaactccct ggttctgttt gaaggggacg
 20 601 aggtgtattc caccatccgg aagcaggaat acaatgggaa gatccctcgg
 651 ttccgcccga tccggggcga gagtgaactg tacaccagtg atactgtcat
 701 gcagaaccca cagttcatca aagccaccat cgtgcaccaa gaccaggctt
 25 751 acgatgacaa gatctactac ttcttccgag aggacaatcc tgacaagaat
 801 cctgaggctc ctctcaatgt gtcccggtg gccagttgt gcagggggga
 851 ccaggggtggg gaaagttcac tgtcagtctc caagtggaa acttttctga
 30 901 aagccatgct ggtatgcagt gatgctgcca ccaacaagaa cttcaacagg
 951 ctgcaagacg tcttctgct ccctgacccc agcggccagt ggagggacac
 1001 caggggtctat ggtgttttct ccaaccctg gaactactca gccgtctgtg
 35 1051 tgtattccct cggtgacatt gacaaggctt tccgtacctc ctactcaag
 1101 ggctaccact caagccttcc caaccgcgg cctggcaagt gcctcccaga
 40 1151 ccagcagccg ataccacacag agaccttcca ggtggctgac cgtcaccag
 1201 aggtggcgca gaggggtggag cccatggggc ctctgaagac gccattgttc
 1251 cactctaaat accactacca gaaagtggcc gttcaccgca tgcaagccag
 45 1301 ccacggggag acctttcatg tgctttacct aactacagac aggggcacta
 1351 tccacaaggt ggtggaaccg ggggagcagg agcacagctt cgccttcaac
 1401 atcatggaga tccagccctt ccgccgcgcg gctgccatcc agaccatgtc
 50 1451 gctggatgct gagcggagga agctgtatgt gagctcccag tgggaggtga
 1501 gccaggtgcc cctggacctg tgtgaggctt atggcggggg ctgccacggt

5

10

15

20

25

30

35

40

45

50

55

1551 tgcctcatgt cccgagaccc ctactgcggc tgggaccagg gccgctgcat
 1601 ctccatctac agctccgaac ggtcagtgtc gcaatccatt aatccagccg
 1651 agccacacaa ggagtgtccc aaccccaaac cagacaaggc cccactgcag
 1701 aaggtttccc tggcccaaaa ctctcgctac tacctgagct gccccatgga
 1751 atccccccac gccacctact catggcgcca caaggagaac gtggagcaga
 1801 gctgcgaacc tggtcaccag agccccaact gcatcctgtt catcgagaac
 1851 ctcacggcgc agcagtacgg ccactacttc tgcgaggccc aggagggtc
 1901 ctacttcgcg gaggctcagc actggcagct gctgcccag gacggcatca
 1951 tggccgagca cctgctgggt catgcctgtg ccctggctgc ctgaattc
 GA
 2001 AGCTTGGAGT CGACTCTGCT GAAGAGGAGG AAATTCTCCT TGAAGTTTCC
 2051 CTGGTGTTCA AAGTAAAGGA GTTGCACCA GACGCACCTC TGTTCACTGG
 2101 TCCGGCGTAT TAAAACACGA TACATTGTTA TTAGTACATT TATTAAGCGC
 2151 TAGATTCTGT GCGTTGTTGA TTTACAGACA ATTGTTGTAC GTATTTTAAT
 2201 AATTCATTAA ATTTATAATC TTTAGGGTGG TATGTTAGAG CGAAAATCAA
 2251 ATGATTTTCA GCGTCTTTAT ATCTGAATTT AAATATTAAA TCCTCAATAG
 2301 ATTTGTAAAA TAGGTTTCGA TTAGTTTCAA ACAAGGGTTG TTTTCCGAA
 2351 CCGATGGCTG GACTATCTAA TGGATTTTCG CTCAACGCCA CAAAAC TTGC
 2401 CAAATCTTGT AGCAGCAATC TAGCTTTGTC GATATTCGTT TGTGTTTTGT
 2451 TTTGTAATAA AGGTTTCGACG TCGTTCAAAA TATTATGCGC TTTTGTATTT
 2501 CTTTCATCAC TGTCGTTAGT GTACAATTGA CTCGACGTAA ACACGTAAAA
 2551 TAAAGCCTGG ACATATTTAA CATCGGGCGT GTTAGCTTTA TTAGCCGAT
 2601 TATCGTCGTC GTCCCAACCC TCGTCGTTAG AAGTTGCTTC CGAAGACGAT
 2651 TTTGCCATAG CCACACGACG CCTATTAATT GTGTCGGCTA ACACGTCCGC
 2701 GATCAAATTT GTAGTTGAGC TTTTGGGAAT TATTTCTGAT TGCGGGCGTT
 2751 TTTGGGCGGG TTTCAATCTA ACTGTGCCCCG ATTTTAATTC AGACAACACG
 2801 TTAGAAAGCG ATGGTGCAGG CCGTGGTAAC ATTTAGACG GCAAATCTAC
 2851 TAATGGCGGC GGTGGTGGAG CTGATGATAA ATCTACCATC GGTGGAGGCG
 2901 CAGGCGGGGC TGGCGGCGGA GGCGGAGGCG GAGGTGGTGG CGGTGATGCA
 2951 GACGGCGGTT TAGGCTCAAA TTGTCTCTTT CAGGCAACAC AGTCGGCACC

5
10
15
20
25
30
35
40
45
50
55

3001 TCAACTATTG TACTGGTTTC GGGCGTATGG TGCACTCTCA GTACAATCTG
3051 CTCTGATGCC GCATAGTTAA GCCAGCCCCG ACACCCGCCA ACACCCGCTG
3101 ACGCGCCCTG ACGGGCTTGT CTGCTCCCGG CATCCGCTTA CAGACAAGCT
3151 GTGACCGTCT CCGGGAGCTG CATGTGTCAG AGGTTTTTAC CGTCATCACC
3201 GAAACGCGCG AGACGAAAGG GCCTCGTGAT ACGCCTATTT TTATAGGTAA
3251 ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA
3301 AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT
3351 GTATCCGCTC ATGAGACAAT AACCTGATA AATGCTTCAA TAATATTGAA
3401 AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT TATTCCCTTT
3451 TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA CGCTGGTGAA
3501 AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC
3551 TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT
3601 TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC
3651 CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
3701 AGAATGACTT GGTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT
3751 GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA
3801 CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG AAGGAGCTAA
3851 CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG
3901 GAACCGGAGC TGAATGAAGC CATAACAAAC GACGAGCGTG ACACCACGAT
3951 GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCGAACTAC
4001 TTA CTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GGCGGATAAA
4051 GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGT TTATTGC
4101 TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC ATTGCAGCAC
4151 TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG
4201 AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC
4251 CTC ACTGATT AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC
4301 TTTAGATTGA TTTAAA ACTT CATTTT TAAT TAAAAGGAT CTAGGTGAAG
4351 ATCCTTTTTG ATAATCTCAT GACCAAATC CCTTAACGTG AGTTTTCGTT
4401 CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC
4451 CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA

5

10

15

20

25

30

35

40

45

50

55

4501 CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTCCGAA
 4551 GGTAAGTGGC TTCAGCAGAG CGCAGATACC AAATACTGTT CTTCTAGTGT
 4601 AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC
 4651 CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC
 4701 GTGTCTTACC GGGTTGGA CTCAAGACGATA GTTACCGGAT AAGGCGCAGC
 4751 GGTCGGGCTG AACGGGGGGT TCGTGACAC AGCCAGCTT GGAGCGAACG
 4801 ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC
 4851 GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG
 4901 GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT
 4951 TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG
 5001 ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGGCCT
 5051 TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTTCCT
 5101 GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC
 5151 TGATACCGCT CGCCGACGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG
 5201 AGGAAGCATC CTGCACCATC GTCTGCTCAT CCATGACCTG ACCATGCAGA
 5251 GGATGATGCT CGTGACGGTT AACGCCTCGA ATCAGCAACG GCTTGCCGTT
 5301 CAGCAGCAGC AGACCATTTT CAATCCGCAC CTCGCGGAAA CCGACATCGC
 5351 AGGCTTCTGC TTCAATCAGC GTGCCGTCGG CGGTGTGCAG TTCAACCACC
 5401 GCACGATAGA GATTCGGGAT TTCGGCGCTC CACAGTTTCG GGTTTTCGAC
 5451 GTTCAGACGT AGTGTGACGC GATCGGTATA ACCACCACGC TCATCGATAA
 5501 TTTACCGGCC GAAAGGCGCG GTGCCGCTGG CGACCTGCGT TTCACCCTGC
 5551 CATAAAGAAA CTGTTACCCG TAGGTAGTCA CGCAACTCGC CGCACATCTG
 5601 AACTTCAGCC TCCAGTACAG CGCGGCTGAA ATCATCATTA AAGCGAGTGG
 5651 CAACATGGAA ATCGCTGATT TGTGTAGTCG GTTTATGCAG CAACGAGACG
 5701 TCACGGAAAA TGCCGCTCAT CCGCCACATA TCCTGATCTT CCAGATAACT
 5751 GCCGTCAC TC AACGCAGCA CCATCACC GC GAGGCGGTTT TCTCCGGCGC
 5801 GTAAAAATGC GCTCAGGTCA AATTCAGACG GCAAACGACT GTCCTGGCCG
 5851 TAACCGACCC AGCGCCCGTT GCACCACAGA TGAAACGCCG AGTTAACGCC
 5901 ATCAAAAATA ATTCGCGTCT GGCCTTCCTG TAGCCAGCTT TCATCAACAT

5951 TAAATGTGAG CGAGTAACAA CCCGTCGGAT TCTCCGTGGG AACAAACGGC
6001 GGATTGACCG TAATGGGATA GGTCACGTTG GTGTAGATGG GCGCATCGTA
5 6051 ACCGTGCATC TGCCAGTTTG AGGGGACGAC GACAGTATCG GCCTCAGGAA
6101 GATCGCACTC CAGCCAGCTT TCCGGCACCG CTTCTGGTGC CGGAAACCAG
10 6151 GCAAAGCGCC ATTCGCCATT CAGGCTGCGC AACTGTTGGG AAGGGCGATC
6201 GGTGCGGGCC TCTTCGCTAT TACGCCAGCT GGCGAAAGGG GGATGTGCTG
6251 CAAGGCGATT AAGTTGGGTA ACGCCAGGGT TTTCCCAGTC ACGACGTTGT
15 6301 AAAACGACGG GATCTATCAT TTTTAGCAGT GATTCTAATT GCAGCTGCTC
6351 TTTGATACAA CTAATTTTAC GACGACGATG CGAGCTTTTA TTCAACCGAG
20 6401 CGTG CATGTT TGCAATCGTG CAAGCGTTAT CAATTTTCA TTATCGTATT
6451 GTTGACATC AACAGGCTGG ACACCACGTT GAACTCGCCG CAGTTTTGCG
6501 GCAAGTTGGA CCCGCCGCGC ATCCAATGCA AACTTTCCGA CATTCTGTTG
25 6551 CCTACGAACG ATTGATTCTT TGTCCATTGA TCGAAGCGAG TGCCTTCGAC
6601 TTTTTCGTGT CCAGTGTGGC TT

30

35

40

45

50

55

SEQUENZPROTOKOLL

(1) ALLGEMEINE INFORMATION:

(i) ANMELDER:

(A) NAME: Hoechst Marion Roussel Deutschland GmbH
(B) STRASSE: -
(C) ORT: Frankfurt
(D) BUNDESLAND: -
(E) LAND: Deutschland
(F) POSTLEITZAHL: 65926
(G) TELEPHON: 069-305-7072
(H) TELEFAX: 069-35-7175
(I) TELEX: -

(ii) ANMELDETITEL: Humanes Semaphorin L (H-Sema-L) und
korrespondierende Semaphorine in anderen Spezies

(iii) ANZAHL DER SEQUENZEN: 44

(iv) COMPUTER-LESBARE FORM:

(A) DATENTRÄGER: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)

(2) INFORMATION ZU SEQ ID NO: 1:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 2636 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..2636

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| CGGGGCCACG | GGATGACGCC | TCCTCCGCCC | GGACGTGCCG | CCCCCAGCGC | ACCGCGCGCC | 60 |
| CGCGTCCCTG | GCCCGCCGGC | TCGGTTGGGG | CTTCCGCTGC | GGCTGCGGCT | GCTGCTGCTG | 120 |
| CTCTGGGCGG | CCGCCGCCTC | CGCCCAGGGC | CACCTAAGGA | GCGGACCCCG | CATCTTCGCC | 180 |
| GTCTGGAAAG | GCCATGTAGG | GCAGGACCGG | GTGGACTTTG | GCCAGACTGA | GCCGCACACG | 240 |
| GTGCTTTTCC | ACGAGCCAGG | CAGCTCCTCT | GTGTGGGTGG | GAGGACGTGG | CAAGGTCTAC | 300 |

CTCTTTGACT TCCCCGAGGG CAAGAACGCA TCTGTGCGCA CGGTGAATAT CGGCTCCACA 360
 AAGGGGTCCT GTCTGGATAA GCGGGACTGC GAGAACTACA TCACTCTCCT GGAGAGGCGG 420
 5 AGTGAGGGGC TGCTGGCCTG TGGCACCAAC GCCCCGCACC CCAGCTGCTG GAACCTGGTG 480
 AATGGCACTG TGGTGCCACT TGGCGAGATG AGAGGCTACG CCCCCTTCAG CCCGGACGAG 540
 10 AACTCCCTGG TTCTGTTTGA AGGGGACGAG GTGTATTCCA CCATCCGGA GCAGGAATAC 600
 AATGGGAAGA TCCCTCGGTT CCGCCGCATC CGGGGCGAGA GTGAGCTGTA CACCAGTGAT 660
 ACTGTCATGC AGAACCCACA GTTCATCAAA GCCACCATCG TGCACCAAGA CCAGGCTTAC 720
 15 GATGACAAGA TCTACTACTT CTTCCGAGAG GACAATCCTG ACAAGAATCC TGAGGCTCCT 780
 CTCAATGTGT CCCGTGTGGC CCAGTTGTGC AGGGGGGACC AGGGTGGGGA AAGTTCACTG 840
 TCAGTCTCCA AGTGGAACAC TTTTCTGAAA GCCATGCTGG TATGCAGTGA TGCTGCCACC 900
 20 AACAAGAACT TCAACAGGCT GCAAGACGTC TTCTGTCTCC CTGACCCCAG CGGCCAGTGG 960
 AGGGACACCA GGGTCTATGG TGTTTTCTCC AAGCCCTGGA ACTACTCAGC CGTCTGTGTG 1020
 TATTCCTCG GTGACATTGA CAAGGTCTTC CGTACCTCCT CACTCAAGGG CTACCACTCA 1080
 25 AGCCTTCCCA ACCCGCGGCC TGGCAAGTGC CTCCCAGACC AGCAGCCGAT ACCCACAGAG 1140
 ACCTTCCAGG TGGCTGACCG TCACCCAGAG GTGGCGCAGA GGGTGGAGCC CATGGGGCCT 1200
 30 CTGAAGACGC CATTGTTCCA CTCTAAATAC CACTACCAGA AAGTGGCCGT TCACCGCATG 1260
 CAAGCCAGCC ACGGGGAGAC CTTTCATGTG CTTTACCTAA CTACAGACAG GGGCACTATC 1320
 CACAAGGTGG TGAACCGGG GGAGCAGGAG CACAGCTTCG CCTTCAACAT CATGGAGATC 1380
 35 CAGCCCTTCC GCCGCGCGGC TGCCATCCAG ACCATGTGCG TGGATGCTGA GCGGAGGAAG 1440
 CTGTATGTGA GCTCCCAGTG GGAGGTGAGC CAGGTGCCCC TGGACCTGTG TGAGGTCTAT 1500
 GGCGGGGGCT GCCACGGTTG CCTCATGTCC CGAGACCCCT ACTGCGGCTG GGACCAGGGC 1560
 40 CGCTGCATCT CCATCTACAG CTCCGAACGG TCAGTGCTGC AATCCATTAA TCCAGCCGAG 1620
 CCACACAAGG AGTGTCCCAA CCCCAAACCA GACAAGGCCC CACTGCAGAA GGTTTCCCTG 1680
 GCCCCAAACT CTCGCTACTA CCTGAGCTGC CCCATGGAAT CCCGCCACGC CACCTACTCA 1740
 45 TGCGCCACA AGGAGAACGT GGAGCAGAGC TGCGAACCTG GTCACCAGAG CCCCAACTGC 1800
 ATCCTGTTCA TCGAGAACCT CACGGCGCAG CAGTACGGCC ACTACTTCTG CGAGGCCAG 1860
 50 GAGGGCTCCT ACTTCCGCGA GGCTCAGCAC TGGCAGCTGC TGCCCGAGGA CGGCATCATG 1920
 GCCGAGCACC TGCTGGGTCA TGCCTGTGCC CTGGCTGCCT CCCTCTGGCT GGGGGTGCTG 1980

CCCACACTCA CTCTTGGCTT GCTGGTCCAC TAGGGCCTCC CGAGGCTGGG CATGCCTCAG 2040
 GCTTCTGCAG CCCAGGGCAC TAGAACGTCT CACACTCAGA GCCGGCTGGC CCGGGAGCTC 2100
 5 CTTGCCTGCC ACTTCTTCCA GGGGACAGAA TAACCCAGTG GAGGATGCCA GGCCTGGAGA 2160
 CGTCCAGCCG CAGGCGGCTG CTGGGCCCCA GGTGGCGCAC GGATGGTGAG GGGCTGAGAA 2220
 10 TGAGGGCACC GACTGTGAAG CTGGGGCATC GATGACCCAA GACTTTATCT TCTGGAAAAT 2280
 ATTTTTCAGA CTCCTCAAAC TTGACTAAAT GCAGCGATGC TCCCAGCCCA AGAGCCCATG 2340
 GGTCCGGGAG TGGGTTTGA TAGGAGAGCT GGGACTCCAT CTCGACCCTG GGGCTGAGGC 2400
 15 CTGAGTCCTT CTGGACTCTT GGTACCCACA TTGCCTCCTT CCCCTCCCTC TCTCATGGCT 2460
 GGGTGGCTGG TGTTCTGAA GACCCAGGGC TACCCTCTGT CCAGCCCTGT CCTCTGCAGC 2520
 TCCCTCTCTG GTCCTGGGTC CCACAGGACA GCCGCCTTGC ATGTTTATTG AAGGATGTTT 2580
 20 GCTTTCCGGA CGGAAGGACG GAAAAAGCTC TGAAAAA AAAA AAAA 2636

(2) INFORMATION ZU SEQ ID NO: 2:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 1195 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
 (B) LAGE: 1..1195

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

CGGGGCTGCG GGATGACGCC TCCTCCTCCC GGACGTGCCG CCCCCAGCGC ACCGCGCGCC 60
 CGCGTCCTCA GCCTGCCGGC TCGGTTTCGGG CTCCCGCTGC GGCTGCGGCT TCTGCTGGTG 120
 45 TTCTGGGTGG CCGCCGCCTC CGCCCAAGGC CACTCGAGGA GCGGACCCCG CATCTCCGCC 180
 GTCTGGAAG GGCAGGACCA TGTGGACTTT AGCCAGCCTG AGCCACACAC CGTGCTTTTC 240
 CATGAGCCGG GCAGCTTCTC TGTCTGGGTG GGTGGACGTG GCAAGGTCTA CCACTTCAAC 300
 50 TTCCCCGAGG GCAAGAATGC CTCTGTGCGC ACGGTGAACA TCGGCTCCAC AAAGGGGTCC 360
 TGTCAGGACA AACAGGACTG TGGGAATTAC ATCACTCTTC TAGAAAGGCG GGGTAATGGG 420

CTGCTGGTCT GTGGCACCAA TGCCCGGAAG CCCAGCTGCT GGAAGTTGGT GAATGACAGT 480
 GTGGTGATGT CACTTGGTGA GATGAAAGGC TATGCCCCCT TCAGCCCGGA TGAGAACTCC 540
 5 CTGGTTCTGT TTGAAGGAGA TGAAGTGAC TCTACCATCC GGAAGCAGGA ATACAACGGG 600
 AAGATCCCTC GGTTTCGACG CATTGCGGGC GAGAGTGAAC TGTACACAAG TGATACAGTC 660
 10 ATGCAGAACC CACAGTTCAT CAAGGCCACC ATTGTGCACC AAGACCAAGC CTATGATGAT 720
 AAGATCTACT ACTTCTTCCG AGAAGACAAC CCTGACAAGA ACCCCGAGGC TCCTCTCAAT 780
 GTGTCCCAG TAGCCAGTT GTGCAGGGGG GACCAGGGTG GTGAGAGTTC GTTGTCTGTC 840
 15 TCCAAGTGA ACACCTTCCT GAAAGCCATG TTGGTCTGCA GCGATGCAGC CACCAACAGG 900
 AACTTCAATC GGCTGCAAGA TGTCTTCCTG CTCCCTGACC CCAGTGGCCA GTGGAGAGAT 960
 20 ACCAGGTCT ATGGCGTTT CTCCAACCCC TGGAACTACT CAGCTGTCTG CGTGTATTCTG 1020
 CTTGGTGACA TTGACAGAGT CTTCCGTACC TCATCGCTCA AAGGCTACCA CATGGGCCTT 1080
 TCCAACCCTC GACCTGGCAT GTGCCTCCCA AAAAAGCAGC CCATACCCAC AGAAACCTTC 1140
 25 CAGGTAGCTG ATAGTCACCC AGAGGTGGCT CAGAGGGTGG AACCTATGGG GCCCC 1195

(2) INFORMATION ZU SEQ ID NO: 3:

30 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 666 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzel
 35 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:
 40 (A) NAME/SCHLÜSSEL: Protein
 (B) LÄGE: 1..666

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

45 Met Thr Pro Pro Pro Pro Gly Arg Ala Ala Pro Ser Ala Pro Arg Ala
 1 5 10 15
 Arg Val Pro Gly Pro Pro Ala Arg Leu Gly Leu Pro Leu Arg Leu Arg
 20 25 30
 50 Leu Leu Leu Leu Trp Ala Ala Ala Ala Ser Ala Gln Gly His Leu
 35 40 45

55

Arg Ser Gly Pro Arg Ile Phe Ala Val Trp Lys Gly His Val Gly Gln
 50 55 60
 5 Asp Arg Val Asp Phe Gly Gln Thr Glu Pro His Thr Val Leu Phe His
 65 70 75 80
 Glu Pro Gly Ser Ser Ser Val Trp Val Gly Gly Arg Gly Lys Val Tyr
 85 90 95
 10 Leu Phe Asp Phe Pro Glu Gly Lys Asn Ala Ser Val Arg Thr Val Asn
 100 105 110
 Ile Gly Ser Thr Lys Gly Ser Cys Leu Asp Lys Arg Asp Cys Glu Asn
 115 120 125
 15 Tyr Ile Thr Leu Leu Glu Arg Arg Ser Glu Gly Leu Leu Ala Cys Gly
 130 135 140
 20 Thr Asn Ala Arg His Pro Ser Cys Trp Asn Leu Val Asn Gly Thr Val
 145 150 155 160
 Val Pro Leu Gly Glu Met Arg Gly Tyr Ala Pro Phe Ser Pro Asp Glu
 165 170 175
 25 Asn Ser Leu Val Leu Phe Glu Gly Asp Glu Val Tyr Ser Thr Ile Arg
 180 185 190
 Lys Gln Glu Tyr Asn Gly Lys Ile Pro Arg Phe Arg Arg Ile Arg Gly
 195 200 205
 30 Glu Ser Glu Leu Tyr Thr Ser Asp Thr Val Met Gln Asn Pro Gln Phe
 210 215 220
 Ile Lys Ala Thr Ile Val His Gln Asp Gln Ala Tyr Asp Asp Lys Ile
 225 230 235 240
 Tyr Tyr Phe Phe Arg Glu Asp Asn Pro Asp Lys Asn Pro Glu Ala Pro
 245 250 255
 40 Leu Asn Val Ser Arg Val Ala Gln Leu Cys Arg Gly Asp Gln Gly Gly
 260 265 270
 Glu Ser Ser Leu Ser Val Ser Lys Trp Asn Thr Phe Leu Lys Ala Met
 275 280 285
 45 Leu Val Cys Ser Asp Ala Ala Thr Asn Lys Asn Phe Asn Arg Leu Gln
 290 295 300
 Asp Val Phe Leu Leu Pro Asp Pro Ser Gly Gln Trp Arg Asp Thr Arg
 305 310 315 320
 50 Val Tyr Gly Val Phe Ser Asn Pro Trp Asn Tyr Ser Ala Val Cys Val
 325 330 335

Tyr Ser Leu Gly Asp Ile Asp Lys Val Phe Arg Thr Ser Ser Leu Lys
 340 345 350
 5 Gly Tyr His Ser Ser Leu Pro Asn Pro Arg Pro Gly Lys Cys Leu Pro
 355 360 365
 Asp Gln Gln Pro Ile Pro Thr Glu Thr Phe Gln Val Ala Asp Arg His
 370 375 380
 10 Pro Glu Val Ala Gln Arg Val Glu Pro Met Gly Pro Leu Lys Thr Pro
 385 390 395 400
 Leu Phe His Ser Lys Tyr His Tyr Gln Lys Val Ala Val His Arg Met
 15 405 410 415
 Gln Ala Ser His Gly Glu Thr Phe His Val Leu Tyr Leu Thr Thr Asp
 420 425 430
 20 Arg Gly Thr Ile His Lys Val Val Glu Pro Gly Glu Gln Glu His Ser
 435 440 445
 Phe Ala Phe Asn Ile Met Glu Ile Gln Pro Phe Arg Arg Ala Ala Ala
 450 455 460
 25 Ile Gln Thr Met Ser Leu Asp Ala Glu Arg Arg Lys Leu Tyr Val Ser
 465 470 475 480
 Ser Gln Trp Glu Val Ser Gln Val Pro Leu Asp Leu Cys Glu Val Tyr
 30 485 490 495
 Gly Gly Gly Cys His Gly Cys Leu Met Ser Arg Asp Pro Tyr Cys Gly
 500 505 510
 35 Trp Asp Gln Gly Arg Cys Ile Ser Ile Tyr Ser Ser Glu Arg Ser Val
 515 520 525
 Leu Gln Ser Ile Asn Pro Ala Glu Pro His Lys Glu Cys Pro Asn Pro
 530 535 540
 40 Lys Pro Asp Lys Ala Pro Leu Gln Lys Val Ser Leu Ala Pro Asn Ser
 545 550 555 560
 Arg Tyr Tyr Leu Ser Cys Pro Met Glu Ser Arg His Ala Thr Tyr Ser
 565 570 575
 45 Trp Arg His Lys Glu Asn Val Glu Gln Ser Cys Glu Pro Gly His Gln
 580 585 590
 Ser Pro Asn Cys Ile Leu Phe Ile Glu Asn Leu Thr Ala Gln Gln Tyr
 595 600 605
 50 Gly His Tyr Phe Cys Glu Ala Gln Glu Gly Ser Tyr Phe Arg Glu Ala
 610 615 620
 55

Gln His Trp Gln Leu Leu Pro Glu Asp Gly Ile Met Ala Glu His Leu
625 630 635 640

Leu Gly His Ala Cys Ala Leu Ala Ala Ser Leu Trp Leu Gly Val Leu
645 650 655

Pro Thr Leu Thr Leu Gly Leu Leu Val His
660 665

(2) INFORMATION ZU SEQ ID NO: 4:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 394 Aminosäuren
(B) ART: Aminosäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: Protein
(B) LÄNGE: 1..394

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

Met Thr Pro Pro Pro Pro Gly Arg Ala Ala Pro Ser Ala Pro Arg Ala
1 5 10 15

Arg Val Leu Ser Leu Pro Ala Arg Phe Gly Leu Pro Leu Arg Leu Arg
20 25 30

Leu Leu Leu Val Phe Trp Val Ala Ala Ser Ala Gln Gly His Ser
35 40 45

Arg Ser Gly Pro Arg Ile Ser Ala Val Trp Lys Gly Gln Asp His Val
50 55 60

Asp Phe Ser Gln Pro Glu Pro His Thr Val Leu Phe His Glu Pro Gly
65 70 75 80

Ser Phe Ser Val Trp Val Gly Gly Arg Gly Lys Val Tyr His Phe Asn
85 90 95

Phe Pro Glu Gly Lys Asn Ala Ser Val Arg Thr Val Asn Ile Gly Ser
100 105 110

Thr Lys Gly Ser Cys Gln Asp Lys Gln Asp Cys Gly Asn Tyr Ile Thr
115 120 125

Leu Leu Glu Arg Arg Gly Asn Gly Leu Leu Val Cys Gly Thr Asn Ala
130 135 140

Arg Lys Pro Ser Cys Trp Asn Leu Val Asn Asp Ser Val Val Met Ser
 145 150 155 160
 5 Leu Gly Glu Met Lys Gly Tyr Ala Pro Phe Ser Pro Asp Glu Asn Ser
 165 170 175
 Leu Val Leu Phe Glu Gly Asp Glu Val Tyr Ser Thr Ile Arg Lys Gln
 180 185 190
 10 Glu Tyr Asn Gly Lys Ile Pro Arg Phe Arg Arg Ile Arg Gly Glu Ser
 195 200 205
 Glu Leu Tyr Thr Ser Asp Thr Val Met Gln Asn Pro Gln Phe Ile Lys
 210 215 220
 15 Ala Thr Ile Val His Gln Asp Gln Ala Tyr Asp Asp Lys Ile Tyr Tyr
 225 230 235 240
 20 Phe Phe Arg Glu Asp Asn Pro Asp Lys Asn Pro Glu Ala Pro Leu Asn
 245 250 255
 Val Ser Arg Val Ala Gln Leu Cys Arg Gly Asp Gln Gly Gly Glu Ser
 260 265 270
 25 Ser Leu Ser Val Ser Lys Trp Asn Thr Phe Leu Lys Ala Met Leu Val
 275 280 285
 30 Cys Ser Asp Ala Ala Thr Asn Arg Asn Phe Asn Arg Leu Gln Asp Val
 290 295 300
 Phe Leu Leu Pro Asp Pro Ser Gly Gln Trp Arg Asp Thr Arg Val Tyr
 305 310 315 320
 35 Gly Val Phe Ser Asn Pro Trp Asn Tyr Ser Ala Val Cys Val Tyr Ser
 325 330 335
 Leu Gly Asp Ile Asp Arg Val Phe Arg Thr Ser Ser Leu Lys Gly Tyr
 340 345 350
 40 His Met Gly Leu Ser Asn Pro Arg Pro Gly Met Cys Leu Pro Lys Lys
 355 360 365
 45 Gln Pro Ile Pro Thr Glu Thr Phe Gln Val Ala Asp Ser His Pro Glu
 370 375 380
 Val Ala Gln Arg Val Glu Pro Met Gly Pro
 385 390
 50
 55

(2) INFORMATION ZU SEQ ID NO: 5:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 23 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..23

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

ACTCACTATA GGGCTCGAGC GGC

23

(2) INFORMATION ZU SEQ ID NO: 6:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 20 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

AGCCGCACAC GGTGCTTTTC

20

(2) INFORMATION ZU SEQ ID NO: 7:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 20 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

GCACAGATGC GTTCTTGCCC

20

5

(2) INFORMATION ZU SEQ ID NO: 8:

(i) SEQUENZ CHARAKTERISTIKA:

10

(A) LÄNGE: 20 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

15

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon

(B) LAGE: 1..20

20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

ACCATAGACC CTGGTGTCCC

20

25

(2) INFORMATION ZU SEQ ID NO: 9:

(i) SEQUENZ CHARAKTERISTIKA:

30

(A) LÄNGE: 20 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

35

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon

(B) LAGE: 1..20

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

GCAGTGATGC TGCCACCAAC

20

45

(2) INFORMATION ZU SEQ ID NO: 10:

(i) SEQUENZ CHARAKTERISTIKA:

50

(A) LÄNGE: 20 Basenpaare

(B) ART: Nukleinsäure

55

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

5 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon

10 (B) LAGE: 1..20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

15 CCAGACCATG TCGCTGGATG

20

(2) INFORMATION ZU SEQ ID NO: 11:

20 (i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 20 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

30 (A) NAME/SCHLÜSSEL: exon

(B) LAGE: 1..20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

35 ACATGAGGCA ACCGTGGCAG

20

(2) INFORMATION ZU SEQ ID NO: 12:

40 (i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 27 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

45 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

50 CCATCCTAAT ACGACTCACT ATAGGGC

27

55

(2) INFORMATION ZU SEQ ID NO: 13:

- 5 (i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 20 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear
- 10 (ii) ART DES MOLEKÜLS: DNS (genomisch)
- (ix) MERKMALE:
15 (A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..20
- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:
20 AGGTAGACCT TGCCACGTCC 20

(2) INFORMATION ZU SEQ ID NO: 14:

- 25 (i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 23 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear
- 30 (ii) ART DES MOLEKÜLS: DNS (genomisch)
- (ix) MERKMALE:
35 (A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..23
- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:
40 GAACTTCAAC AGGCTGCAAG ACG 23

(2) INFORMATION ZU SEQ ID NO: 15:

- 45 (i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 20 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear
- 50 (ii) ART DES MOLEKÜLS: DNS (genomisch)
- 55

- (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LÄGE: 1..20

5

- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 15:

ATGCTGAGCG GAGGAAGCTG

20

10

- (2) INFORMATION ZU SEQ ID NO: 16:

- (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

15

- (ii) ART DES MOLEKÜLS: DNS (genomisch)

20

- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 16:

CCGCCATACA CCTCACACAG

20

25

- (2) INFORMATION ZU SEQ ID NO: 17:

- (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 28 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

30

- (ii) ART DES MOLEKÜLS: DNS (genomisch)

35

- (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LÄGE: 1..28

40

- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 17:

CTGGAAGCTT TCTGTGGGTA TCGGCTGC

28

45

- (2) INFORMATION ZU SEQ ID NO: 18:

- (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 25 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

50

55

(ii) ART DES MOLEKÜLS: DNS (genomisch)

5

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon

(B) LAGE: 1..25

10

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 18:

TTTGGATCCC TGGTTCTGTT TGAAG

25

15

(2) INFORMATION ZU SEQ ID NO: 19:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 50 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

20

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

25

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon

(B) LAGE: 1..50

30

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 19:

TTCTAGAATT CAGCGCCCGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT

50

35

(2) INFORMATION ZU SEQ ID NO: 20:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 27 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

40

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

45

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon

(B) LAGE: 1..27

50

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 20:

GGGGAAAGTT CACTGTCAGT CTCCAAG

27

55

(2) INFORMATION ZU SEQ ID NO: 21:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 26 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 21:

GGGAATACAC ACAGACGGCT GAGTAG

26

(2) INFORMATION ZU SEQ ID NO: 22:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 22 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..22

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 22:

AGCAAGTTCA GCCTGGTTAA GT

22

(2) INFORMATION ZU SEQ ID NO: 23:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 21 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..21

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 23:

TTATGAGTAT TTCTTCCAGG G

21

5

(2) INFORMATION ZU SEQ ID NO: 24:

(i) SEQUENZ CHARAKTERISTIKA:

10

- (A) LÄNGE: 26 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

15

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..26

20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 24:

CCATTAATCC AGCCGAGCCA CACAAG

26

25

(2) INFORMATION ZU SEQ ID NO: 25:

(i) SEQUENZ CHARAKTERISTIKA:

30

- (A) LÄNGE: 25 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

35

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..25

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 25:

CATCTACAGC TCCGAACGGT CAGTG

25

45

(2) INFORMATION ZU SEQ ID NO: 26:

(i) SEQUENZ CHARAKTERISTIKA:

50

- (A) LÄNGE: 20 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

55

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..20

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 26:

CAGCGGAAGC CCCAACCGAG

20

(2) INFORMATION ZU SEQ ID NO: 27:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 23 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..23

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 27:

GGGATGACGC CTCCTCCGCC CGG

23

(2) INFORMATION ZU SEQ ID NO: 28:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 31 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..31

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 28:

AAGCTTCACG TGGACCAGCA AGCCAAGAGT G

31

(2) INFORMATION ZU SEQ ID NO: 29:

(i) SEQUENZ CHARAKTERISTIKA:

- 5 (A) LÄNGE: 25 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- 15 (A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 29:

20 AAGCTTTTTTC CGTCCTTCCG TCCGG

25

(2) INFORMATION ZU SEQ ID NO: 30:

(i) SEQUENZ CHARAKTERISTIKA:

- 25 (A) LÄNGE: 24 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

30 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- 35 (A) NAME/SCHLÜSSEL: exon
(B) LAGE: 1..24

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 30:

40 ATGGTGAGCA AGGCGGAGGA GCTG

24

(2) INFORMATION ZU SEQ ID NO: 31:

(i) SEQUENZ CHARAKTERISTIKA:

- 45 (A) LÄNGE: 24 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

50 (ii) ART DES MOLEKÜLS: DNS (genomisch)

55

- (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LÄGE: 1..24

5

- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 31:

CTTGTACAGC TCGTCCATGC CGAG

24

10

- (2) INFORMATION ZU SEQ ID NO: 32:

- (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 25 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

15

- (ii) ART DES MOLEKÜLS: DNS (genomisch)

20

- (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LÄGE: 1..25

25

- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 32:

GGGTGGTGAG AGTTCGTTGT CTGTC

25

30

- (2) INFORMATION ZU SEQ ID NO: 33:

- (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 25 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

35

- (ii) ART DES MOLEKÜLS: DNS (genomisch)

40

- (ix) MERKMALE:
 (A) NAME/SCHLÜSSEL: exon
 (B) LÄGE: 1..25

45

- (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 33:

GAGCGATGAG GTACGGAAGA CTCTG

25

50

55

(2) INFORMATION ZU SEQ ID NO: 34:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 5856 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
 (B) LAGE: 1..5856

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 34:

| | |
|--|------|
| AGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC | 60 |
| ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC | 120 |
| TCACTCATTA GGCACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA | 180 |
| TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCTTC | 240 |
| ACGTGGACCA GCAAGCCAAG AGTGAGTGTG GGCAGCACCC CCAGCCAGAG GGAGGCAGCC | 300 |
| AGGGCACAGG CATGACCCAG CAGGTGCTCG GCCATGATGC CGTCCTCGGG CAGCAGCTGC | 360 |
| CAGTGCTGAG CCTCGCGGAA GTAGGAGCCC TCCTGGGCCT CGCAGAAGTA GTGGCCGTAC | 420 |
| TGCTGCGCCG TGAGGTTCTC GATGAACAGG ATGCAGTTGG GGCTCTGGTG ACCAGGTTCTG | 480 |
| CAGCTCTGCT CCACGTTCTC CTTGTGGCGC CATGAGTAGG TGGCGTGGCG GGATTCCATG | 540 |
| GGGCAGCTCA GGTAGTAGCG AGAGTTTGGG GCCAGGAAA CCTTCTGCAG TGGGGCCTTG | 600 |
| TCTGGTTTGG GGTGGGACA CTCCTTGTGT GGCTCGGCTG GATTAATGGA TTGCAGCACT | 660 |
| GACCGTTCTG AGCTGTAGAT GGAGATGCAG CGGCCCTGGT CCCAGCCGCA GTAGGGGTCT | 720 |
| CGGGACATGA GGCAACCGTG GCAGCCCCCG CCATAGACCT CACACAGGTC CAGGGGCACC | 780 |
| TGGCTCACCT CCCACTGGGA GCTCACATAC AGCTTCCTCC GCTCAGCATC CAGCGACATG | 840 |
| GTCTGGATGG CAGCCGCGCG GCGGAAGGGC TGGATCTCCA TGATGTTGAA GGCGAAGCTG | 900 |
| TGCTCCTGCT CCCCCGGTTC CACCACCTTG TGGATAGTGC CCCTGTCTGT AGTTAGGTAA | 960 |
| AGCACATGAA AGGTCTCCCC GTGGCTGGCT TGCATGCGGT GAACGGCCAC TTTCTGGTAG | 1020 |
| TGGTATTTAG AGTGAACAA TGGCGTCTTC AGAGGCCCA TGGGCTCCAC CCTCTGCGCC | 1080 |
| ACCTCTGGGT GACGGTCAGC CACCTGGAAG GTCTCTGTGG GTATCGGCTG CTGGTCTGGG | 1140 |

AGGCACTTGC CAGGCCGCGG GTTGGGAAGG CTTGAGTGGT AGCCCTTGAG TGAGGAGGTA 1200
 CGGAAGACCT TGTCAATGTC ACCGAGGGAA TACACACAGA CGGCTGAGTA GTTCCAGGGG 1260
 5 TTGGAGAAAA CACCATAGAC CCTGGTGTCC CTCCACTGGC CGCTGGGGTC AGGGAGCAGG 1320
 AAGACGTCTT GCAGCCTGTT GAAGTTCTTG TTGGTGGCAG CATCACTGCA TACCAGCATG 1380
 10 GCTTTCAGAA AAGTGTTCCA CTTGGAGACT GACAGTGAAC TTTCCCCACC CTGGTCCCCC 1440
 CTGCACAACT GGGCCACACG GGACACATTG AGAGGAGCCT CAGGATTCTT GTCAGGATTG 1500
 TCCTCTCGGA AGAAGTAGTA GATCTTGTCA TCGTAAGCCT GGTCTTGGTG CACGATGGTG 1560
 15 GCTTTGATGA ACTGTGGGTT CTGCATGACA GTATCACTGG TGTACAGCTC ACTCTCGCCC 1620
 CGGATGCGGC GGAACCGAGG GATCTTCCCA TTGTATTCCT GCTTCCGGAT GGTGGAATAC 1680
 ACCTCGTCCC CTTCAAACAG AACCAGGGAG TTCTCGTCCG GGCTGAAGGG GGCGTAGCCT 1740
 20 CTCATCTCGC CAAGTGGCAC CACAGTGCCA TTCACCAGGT TCCAGCAGCT GGGGTGCCGG 1800
 GCGTTGGTGC CACAGGCCAG CAGCCCCCTCA CTCCGCCTCT CCAGGAGAGT GATGTAGTTC 1860
 25 TCGCAGTCCC GCTTATCCAG ACAGGACCCC TTTGTGGAGC CGATATTCAC CGTGCGCACA 1920
 GATGCGTTCT TGCCCTCGGG GAAGTCAAAG AGGTAGACCT TGCCACGTCC TCCCACCCAC 1980
 ACAGAGGAGC TGCCTGGCTC GTGAAAAGC ACCGTGTGCG GCTCAGTCTG GCCAAAGTCC 2040
 30 ACCCGGTCCT GCCCTACATG GCCTTTCCAG ACGGCGAAGA TGCGGGGTCC GCTCCTTAGG 2100
 TGGCCCTGGG CGGAGGCGGC GGCCGCCCAG AGCAGCAGCA GCAGCCGAG CCGCAGCGGA 2160
 35 AGCCCCAACC GAGCCGGCGG GCCAGGGACG CGGGCGCGCG GTGCGCTGGG GGCGGCACGT 2220
 CCGGGCGGAG GAGGCGTCAT CCCAAGCCGA ATTCTGCAGA TATCCATCAC ACTGGCGGCC 2280
 GCTCGAGCAT GCATCTAGAG GGCCCAATTC GCCCTATAGT GAGTCGTATT ACAATTCACT 2340
 40 GGCCGTCGTT TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT 2400
 TGCAGCACAT CCCCCTTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC 2460
 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAAG 2520
 45 CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC 2580
 CGCTCCTTTC GCTTCTTCC CTTCTTTCT CGCCACGTTT GCCGGCTTTC CCCGTCAAGC 2640
 50 TCTAAATCGG GGGCTCCCTT TAGGGTTCCG ATTTAGAGCT TTACGGCACC TCGACCGCAA 2700
 AAAACTTGAT TTGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTTCG 2760

| | | | |
|----|--|---------------------------|------|
| | CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGA | CTC TTGTTCCAAA CTGGAACAAC | 2820 |
| 5 | ACTCAACCCT ATCGCGGTCT ATTCTTTTGA TTTATAAGGG | ATTTTGCCGA TTTCGGCCTA | 2880 |
| | TTGGTTAAAA AATGAGCTGA TTTAACAAAT TCAGGGCGCA | AGGGCTGCTA AAGGAACCGG | 2940 |
| | AACACGTA GA AAGCCAGTCC GCAGAAACGG TGCTGACCCC | GGATGAATGT CAGCTACTGG | 3000 |
| 10 | GCTATCTGGA CAAGGGAAAA CGCAAGCGCA AAGAGAAAGC | AGGTAGCTTG CAGTGGGCTT | 3060 |
| | ACATGGCGAT AGCTAGACTG GGCGGTTTTA TGGACAGCAA | GCGAACC | 3120 |
| | GGGCGCCCT CTGGTAAGGT TGGGAAGCCC TGCAAAGTAA | ACTGGATGGC TTTCTTGCCG | 3180 |
| 15 | CCAAGGATCT GATGGCGCAG GGGATCAAGA TCTGATCAAG | AGACAGGATG AGGATCGTTT | 3240 |
| | CGCATGATTG AACAAGATGG ATTGCACGCA GGTTCTCCGG | CCGCTTGGGT GGAGAGGCTA | 3300 |
| | TTCCGGCTATG ACTGGGCACA ACAGACAATC GGCTGCTCTG | ATGCCGCCGT GTTCCGGCTG | 3360 |
| 20 | TCAGCGCAGG GGCGCCCGGT TCTTTTTGTC AAGACCGACC | TGTCCGGTGC CCTGAATGAA | 3420 |
| | CTGCAGGACG AGGCAGCGCG GCTATCGTGG CTGGCCACGA | CGGGCGTTCC TTGCGCAGCT | 3480 |
| | GTGCTCGACG TTGTCACTGA AGCGGGAAGG GACTGGCTGC | TATTGGGCGA AGTGCCGGGG | 3540 |
| 25 | CAGGATCTCC TGTCACTCG CCTTGCTCCT GCCGAGAAAG | TATCCATCAT GGCTGATGCA | 3600 |
| | ATGCGGCGGC TGCATACGCT TGATCCGGCT ACCTGCCCAT | TCGACCACCA AGCGAAACAT | 3660 |
| 30 | CGCATCGAGC GAGCACGTAC TCGGATGGAA GCCGGTCTTG | TCGATCAGGA TGATCTGGAC | 3720 |
| | GAAGAGCATC AGGGGCTCGC GCCAGCCGAA CTGTTCCGA | GGCTCAAGGC GCGCATGCCC | 3780 |
| | GACGGCGAGG ATCTCGTCGT GATCCATGGC GATGCCTGCT | TGCCGAATAT CATGGTGGAA | 3840 |
| 35 | AATGGCCGCT TTTCTGGATT CAACGACTGT GGCCGGCTGG | GTGTGGCGGA CCGCTATCAG | 3900 |
| | GACATAGCGT TGGATACCCG TGATATTGCT GAAGAGCTTG | GCGGCGAATG GGCTGACCGC | 3960 |
| | TTCTCTGTC TTTACGGTAT CGCCGCTCCC GATTTCGAGC | GCATCGCCTT CTATCGCCTT | 4020 |
| 40 | CTTGACGAGT TCTTCTGAAT TGAAAAAGGA AGAGTATGAG | TATTCAACAT TTCCGTGTCG | 4080 |
| | CCCTTATTCC CTTTTTTGCG GCATTTTGCC TTCCTGTTTT | TGCTCACCCA GAAACGCTGG | 4140 |
| 45 | TGAAAGTAAA AGATGCTGAA GATCAGTTGG GTGCACGAGT | GGGTTACATC GAACTGGATC | 4200 |
| | TCAACAGCGG TAAGATCCTT GAGAGTTTTC GCCCCGAAGA | ACGTTTTCCA ATGATGAGCA | 4260 |
| | CTTTTAAAGT TCTGCTATGT CATACTAT TATCCCGTAT | TGACGCCGGG CAAGAGCAAC | 4320 |
| 50 | TCGGTCGCCG GGCGCGGTAT TCTCAGAATG ACTTGGTTGA | GTA | 4380 |
| | AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG | TGCTGCCATA ACCATGAGTG | 4440 |

| | | |
|----|---|------|
| | ATAACACTGC GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT | 4500 |
| | TTTTGCACAA CATGGGGGAT CATGTAATC GCCTTGATCG TTGGGAACCG GAGCTGAATG | 4560 |
| 5 | AAGCCATACC AAACGACGAG AGTGACACCA CGATGCCTGT AGCAATGCCA ACAACGTTGC | 4620 |
| | GCAAACTATT AACTGGCGAA CTACTTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA | 4680 |
| 10 | TGGAGGCGGA TAAAGTTGCA GGACCACTTC TGCCTCGGC CCTTCCGGCT GGCTGGTTTA | 4740 |
| | TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCG TATCATTGCA GCACTGGGGC | 4800 |
| | CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG | 4860 |
| 15 | ATGAACGAAA TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT | 4920 |
| | CAGACCAAGT TTAATCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA | 4980 |
| | GGATCTAGGT GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT | 5040 |
| 20 | CGTTCCACTG AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT | 5100 |
| | TTCTGCGCGT AATCTGCTGC TTGCAAACAA AAAAACCACC GCTACCAGCG GTGGTTTGTT | 5160 |
| 25 | TGCCGGATCA AGAGCTACCA ACTCTTTTTT CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA | 5220 |
| | TACCAAATAC TGTCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG | 5280 |
| | CACCGCCTAC ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA | 5340 |
| 30 | AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG | 5400 |
| | GCTGAACGGG GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA | 5460 |
| 35 | GATACCTACA GCGTGAGCAT TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGACA | 5520 |
| | GGTATCCGGT AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGGAA | 5580 |
| | ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT CTGACTTGAG CGTCGATTTT | 5640 |
| 40 | TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTTAC | 5700 |
| | GGTTCCTGGC CTTTGTGCTG CTTTGTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT | 5760 |
| | CTGTGGATAA CCGTATTACC GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA | 5820 |
| 45 | CCGAGCGCAG CGAGTCAGTG AGCGAGGAAG CGGAAG | 5856 |

(2) INFORMATION ZU SEQ ID NO: 35:

(i) SEQUENZ CHARAKTERISTIKA:

(A) LÄNGE: 7475 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

5 (ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon

10 (B) LAGE: 1..7475

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 35:

15 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCGACTCT CAGTACAATC TGCTCTGATG 60

CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG 120

CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC 180

20 TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT 240

GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA 300

TGGAGTTCGG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC 360

25 CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC 420

ATTGACGTCA ATGGGTGGAC TATTTACGGT AAAC TGCCCA CTTGGCAGTA CATCAAGTGT 480

30 ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT 540

ATGCCCAGTA CATGACCTTA TGGGACTTTC CTA CTGTTGGCA GTACATCTAC GTATTAGTCA 600

TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG 660

35 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC 720

AAAATCAACG GGA CTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG 780

GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA 840

40 CTGCTTACTG GCTTATCGAA ATTAATACGA CTC ACTATAG GGAGACCCAA GCTGGCTAGC 900

GTTTAAACGG GCCCTCTAGA CTCGAGCGGC CGCCACTGTG CTGGATATCT GCAGAATTCG 960

45 GCTTGGGATG ACGCCTCCTC CGCCCGGACG TGCCGCCCCC AGCGCACCGC GCGCCCGCGT 1020

CCCTGGCCCC CGGCTCGGT TGGGGCTTCC GCTGCGGCTG CGGCTGCTGC TGCTGCTCTG 1080

GGCGGCCGCC GCCTCCGCC AGGGCCACCT AAGGAGCGGA CCCCGCATCT TCGCCGTCTG 1140

50 GAAAGGCCAT GTAGGGCAGG ACCGGGTGGA CTTTGGCCAG ACTGAGCCGC ACACGGTGCT 1200

TTCCACGAG CCAGGCAGCT CCTCTGTGTG GGTGGGAGGA CGTGGCAAGG TCTACCTCTT 1260

55

| | | |
|----|--|------|
| | TGACTTCCCC GAGGGCAAGA ACGCATCTGT GCGCACGGTG AATATCGGCT CCACAAAGGG | 1320 |
| | GTCCTGTCTG GATAAGCGGG ACTGCGAGAA CTACATCACT CTCCTGGAGA GGCGGAGTGA | 1380 |
| 5 | GGGGCTGCTG GCCTGTGGCA CCAACGCCCG GCACCCAGC TGCTGGAACC TGGTGAATGG | 1440 |
| | CACTGTGGTG CCACTTGGCG AGATGAGAGG CTACGCCCCC TTCAGCCCGG ACGAGAACTC | 1500 |
| 10 | CCTGGTTCTG TTTGAAGGGG ACGAGGTGTA TTCCACCATC CGGAAGCAGG AATACAATGG | 1560 |
| | GAAGATCCCT CGGTTCCGCC GCATCCGGGG CGAGAGTGAG CTGTACACCA GTGATACTGT | 1620 |
| | CATGCAGAAC CCACAGTTCA TCAAAGCCAC CATCGTGAC CAAGACCAGG CTTACGATGA | 1680 |
| 15 | CAAGATCTAC TACTTCTTCC GAGAGGACAA TCCTGACAAG AATCCTGAGG CTCCTCTCAA | 1740 |
| | TGTGTCCCGT GTGGCCAGT TGTGCAGGGG GGACCAGGGT GGGGAAAGTT CACTGTCACT | 1800 |
| | CTCCAAGTGG AACACTTTTC TGAAAGCCAT GCTGGTATGC AGTGATGCTG CCACCAACAA | 1860 |
| 20 | GAACTTCAAC AGGCTGCAAG ACGTCTTCCT GCTCCCTGAC CCCAGCGGCC AGTGGAGGGA | 1920 |
| | CACCAGGGTC TATGGTGTTC TCTCCAACCC CTGGAAGTAC TCAGCCGTCT GTGTGTATTC | 1980 |
| | CCTCGGTGAC ATTGACAAGG TCTTCCGTAC CTCCTCACTC AAGGGCTACC ACTCAAGCCT | 2040 |
| 25 | TCCCAACCCG CGGCCTGGCA AGTGCCTCCC AGACCAGCAG CCGATACCCA CAGAGACCTT | 2100 |
| | CCAGGTGGCT GACCGTCACC CAGAGGTGGC GCAGAGGGTG GAGCCCATGG GGCCTCTGAA | 2160 |
| 30 | GACGCCATTG TTCCACTCTA AATACCACTA CCAGAAAGTG GCCGTCACC GCATGCAAGC | 2220 |
| | CAGCCACGGG GAGACCTTTC ATGTGCTTTA CCTAACTACA GACAGGGGCA CTATCCACAA | 2280 |
| | GGTGGTGGAA CCGGGGGAGC AGGAGCACAG CTTGCGCTTC AACATCATGG AGATCCAGCC | 2340 |
| 35 | CTTCCGCCGC GCGGCTGCCA TCCAGACCAT GTCGCTGGAT GCTGAGCGGA GGAAGCTGTA | 2400 |
| | TGTGAGCTCC CAGTGGGAGG TGAGCCAGGT GCCCCTGGAC CTGTGTGAGG TCTATGGCGG | 2460 |
| | GGGCTGCCAC GGTGCTCTCA TGTCCCGAGA CCCCTACTGC GGCTGGGACC AGGGCCGCTG | 2520 |
| 40 | CATCTCCATC TACAGCTCCG AACGGTCAGT GCTGCAATCC ATTAATCCAG CCGAGCCACA | 2580 |
| | CAAGGAGTGT CCCAACCCCA AACCAGACAA GGCCCCACTG CAGAAGGTTT CCCTGGCCCC | 2640 |
| 45 | AAACTCTCGC TACTACCTGA GCTGCCCCAT GGAATCCCGC CACGCCACCT ACTCATGGCG | 2700 |
| | CCACAAGGAG AACGTGGAGC AGAGCTGCGA ACCTGGTCAC CAGAGCCCCA ACTGCATCCT | 2760 |
| | GTTTCATCGAG AACCTCACGG CGCAGCAGTA CGGCCACTAC TTCTGCGAGG CCCAGGAGGG | 2820 |
| 50 | CTCCTACTTC CGCGAGGCTC AGCACTGGCA GCTGCTGCCC GAGGACGGCA TCATGGCCGA | 2880 |
| | GCACCTGCTG GGTGATGCCT GTGCCCTGGC TGCCTCCCTC TGGCTGGGGG TGCTGCCAC | 2940 |

ACTCACTCTT GGCTTGCTGG TCCACGTGAA GCTTGGGCCC GAACAAAAAC TCATCTCAGA 3000
 AGAGGATCTG AATAGCGCCG TCGACCATCA TCATCATCAT CATTGAGTTT AAACCGCTGA 3060
 5 TCAGCCTCGA CTGTGCCTTC TAGTTGCCAG CCATCTGTTG TTTGCCCCCTC CCCCCTGCCT 3120
 TCCTTGACCC TGAAGGTGC CACTCCCACT GTCCTTTCCT AATAAAATGA GGAAATTGCA 3180
 10 TCGCATTGTC TGAGTAGGTG TCATTCTATT CTGGGGGTG GGGTGGGCA GGACAGCAAG 3240
 GGGGAGGATT GGAAGACAA TAGCAGGCAT GCTGGGGATG CGGTGGGCTC TATGGCTTCT 3300
 GAGGCGGAAA GAACCAGCTG GGGCTCTAGG GGGTATCCCC ACGCGCCCTG TAGCGGCGCA 3360
 15 TTAAGCGCGG CGGGTGTGGT GGTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA 3420
 GCGCCCGCTC CTTTCGCTTT CTTCCTTCC TTTCTCGCCA CGTTCGCCGG CTTTCCCCGT 3480
 CAAGCTCTAA ATCGGGGCAT CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC 3540
 20 CCCAAAAAAC TTGATTAGGG TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT 3600
 TTTGCCCCCT TGACGTTGGA GTCCACGTTT TTTAATAGTG GACTCTTGTT CCAAAGTGA 3660
 25 ACAACACTCA ACCCTATCTC GGTCTATTCT TTTGATTAT AAGGGATTTT GGGGATTTCG 3720
 GCCTATTGGT TAAAAAATGA GCTGATTTAA CAAAAATTA ACGCAATTA ATTCTGTGGA 3780
 ATGTGTGTCA GTTAGGGTGT GGAAAGTCCC CAGGCTCCCC AGGCAGGCAG AAGTATGCAA 3840
 30 AGCATGCATC TCAATTAGTC AGCAACCAGG TGTGAAAGT CCCCAGGCTC CCCAGCAGGC 3900
 AGAAGTATGC AAAGCATGCA TCTCAATTAG TCAGCAACCA TAGTCCCGCC CCTAACTCCG 3960
 35 CCCATCCCGC CCCTAACTCC GCCCAGTTCC GCCCATTCTC CGCCCCATGG CTGACTAATT 4020
 TTTTTTATTT ATGCAGAGGC CGAGGCCGCC TCTGCCTCTG AGCTATTCCA GAAGTAGTGA 4080
 GGAGGCTTTT TTGGAGGCCT AGGCTTTTGC AAAAAGCTCC CGGGAGCTTG TATATCCATT 4140
 40 TTCGGATCTG ATCAAGAGAC AGGATGAGGA TCGTTTCGCA TGATTGAACA AGATGGATTG 4200
 CACGCAGGTT CTCCGGCCGC TTGGGTGGAG AGGCTATTCTG GCTATGACTG GGCACAACAG 4260
 ACAATCGGCT GCTCTGATGC CGCCGTGTTT CGGCTGTCAG CGCAGGGGCG CCCGGTTCTT 4320
 45 TTTGTCAAGA CCGACCTGTC CGGTGCCCTG AATGAACTGC AGGACGAGGC AGCGCGGCTA 4380
 TCGTGGCTGG CCACGACGGG CGTTCCTTGC GCAGCTGTGC TCGACGTTGT CACTGAAGCG 4440
 50 GGAAGGGACT GGCTGCTATT GGGCGAAGTG CCGGGGCAGG ATCTCCTGTC ATCTCACCTT 4500
 GCTCCTGCCG AGAAAGTATC CATCATGGCT GATGCAATGC GGCGGCTGCA TACGCTTGAT 4560

| | | |
|----|--|------|
| | CCGGCTACCT GCCCATTCGA CCACCAAGCG AAACATCGCA TCGAGCGAGC ACGTACTCGG | 4620 |
| | ATGGAAGCCG GTCTTGTCTGA TCAGGATGAT CTGGACGAAG AGCATCAGGG GCTCGCGCCA | 4680 |
| 5 | GCCGAACTGT TCGCCAGGCT CAAGGCGCGC ATGCCCGACG GCGAGGATCT CGTCGTGACC | 4740 |
| | CATGGCGATG CCTGCTTGCC GAATATCATG GTGGAAAATG GCCGCTTTTC TGGATTATC | 4800 |
| 10 | GACTGTGGCC GGCTGGGTGT GGCGGACCGC TATCAGGACA TAGCGTTGGC TACCCGTGAT | 4860 |
| | ATTGCTGAAG AGCTTGGCGG CGAATGGGCT GACCGCTTCC TCGTGCTTTA CGGTATCGCC | 4920 |
| | GCTCCCGATT CGCAGCGCAT CGCCTTCTAT CGCCTTCTTG ACGAGTTCTT CTGAGCGGGA | 4980 |
| 15 | CTCTGGGGTT CGAAATGACC GACCAAGCGA CGCCCAACCT GCCATCACGA GATTTTCGATT | 5040 |
| | CCACCGCCGC CTCTATGAA AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA | 5100 |
| | TGATCCTCCA GCGCGGGGAT CTCATGCTGG AGTTCTTCGC CCACCCCAAC TTGTTTATTG | 5160 |
| 20 | CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT AAAGCATTTT | 5220 |
| | TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCTGTA | 5280 |
| 25 | TACCGTCGAC CTCTAGCTAG AGCTTGGCGT AATCATGGTC ATAGCTGTTT CCTGTGTGAA | 5340 |
| | ATTGTTATCC GCTCACAATT CCACACAACA TACGAGCCGG AAGCATAAAG TGTAAGCCCT | 5400 |
| | GGGGTGCCTA ATGAGTGAGC TAACTCACAT TAATTGCGTT GCGTCACTG CCCGCTTTCC | 5460 |
| 30 | AGTCGGGAAA CCTGTCTGTC CAGCTGCATT AATGAATCGG CCAACGCGCG GGGAGAGGCG | 5520 |
| | GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTGTTTC | 5580 |
| | GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG | 5640 |
| 35 | GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA | 5700 |
| | AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCC TGACGAGCAT CACAAAAATC | 5760 |
| | GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC | 5820 |
| 40 | CTGGAAGCTC CCTCGTGCGC TCTCCTGTTT CGACCCTGCC GCTTACCGGA TACCTGTCCG | 5880 |
| | CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCAATGCTC ACGCTGTAGG TATCTCAGTT | 5940 |
| 45 | CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC | 6000 |
| | GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC | 6060 |
| | CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG | 6120 |
| 50 | AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG | 6180 |
| | CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAACAAA | 6240 |
| 55 | | |

CCACCGCTGG TAGCGGTGGT TTTTTTGT TT GCAAGCAGCA GATTACGCGC AGAAAAAAG 6300
 5 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT 6360
 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA 6420
 ATTA AAAATG AAGTTTAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT 6480
 10 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTTCGT TCATCCATAG 6540
 TTGCCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA 6600
 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTATCA GCAATAAACC 6660
 15 AGCCAGCCGG AAGGGCCGAG CGCAGAAAGT GTCCTGCAAC TTTATCCGCC TCCATCCAGT 6720
 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG 6780
 20 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTGGGTATG GCTTCATTCA 6840
 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG 6900
 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGAGTG TTATCACTCA 6960
 25 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG 7020
 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT 7080
 30 CTTGCCCCGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA 7140
 TCATTGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA 7200
 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG 7260
 35 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC 7320
 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT 7380
 40 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC 7440
 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTC 7475

(2) INFORMATION ZU SEQ ID NO: 36:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 8192 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

(A) NAME/SCHLÜSSEL: exon

(B) LAGE: 1..8192

5

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 36:

10

15

20

25

30

35

40

45

50

55

| | |
|--|------|
| GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGACTCT CAGTACAATC TGCTCTGATG | 60 |
| CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG | 120 |
| CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC | 180 |
| TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT | 240 |
| GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA | 300 |
| TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCCGC TGGCTGACCG CCCAACGACC | 360 |
| CCCGCCCATTT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC | 420 |
| ATTGACGTCA ATGGGTGGAC TATTTACGGT AAAGTGGCCA CTTGGCAGTA CATCAAGTGT | 480 |
| ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT | 540 |
| ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA | 600 |
| TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG | 660 |
| ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC | 720 |
| AAAATCAACG GGAAGTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG | 780 |
| GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA | 840 |
| CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC | 900 |
| GTTTAAACGG GCCCTCTAGA CTCGAGCGGC CGCCACTGTG CTGGATATCT GCAGAATTCTG | 960 |
| GCTTGGGATG ACGCCTCCTC CGCCCGGACG TGCCGCCCCC AGCGCACCGC GCGCCCGCGT | 1020 |
| CCCTGGCCCG CCGGCTCGGT TGGGGCTTCC GCTGCGGCTG CGGCTGCTGC TGCTGCTCTG | 1080 |
| GGCGGCCGCC GCCTCCGCC AGGGCCACCT AAGGAGCGGA CCCCGCATCT TCGCCGTCTG | 1140 |
| GAAAGGCCAT GTAGGGCAGG ACCGGGTGGA CTTTGGCCAG ACTGAGCCGC ACACGGTGCT | 1200 |
| TTTCCACGAG CCAGGCAGCT CCTCTGTGTG GGTGGGAGGA CGTGGCAAGG TCTACCTCTT | 1260 |
| TGACTTCCCC GAGGGCAAGA ACGCATCTGT GCGCACGGTG AATATCGGCT CCACAAAGGG | 1320 |
| GTCCTGTCTG GATAAGCGGG ACTGCGAGAA CTACATCACT CTCCTGGAGA GGCGGAGTGA | 1380 |
| GGGGCTGCTG GCCTGTGGCA CCAACGCCCG GCACCCAGC TGCTGGAACC TGGTGAATGG | 1440 |

| | | |
|----|--|------|
| | CACTGTGGTG CCACTTGGCG AGATGAGAGG CTACGCCCCC TTCAGCCCGG ACGAGAACTC | 1500 |
| 5 | CCTGGTTCTG TTTGAAGGGG ACGAGGTGTA TTCCACCATC CGGAAGCAGG AATACAATGG | 1560 |
| | GAAGATCCCT CGGTTCCGCC GCATCCGGGG CGAGAGTGAG CTGTACACCA GTGATACTGT | 1620 |
| | CATGCAGAAC CCACAGTTCA TCAAAGCCAC CATCGTGCAC CAAGACCAGG CTTACGATGA | 1680 |
| 10 | CAAGATCTAC TACTTCTTCC GAGAGGACAA TCCTGACAAG AATCCTGAGG CTCCTCTCAA | 1740 |
| | TGTGTCCCGT GTGGCCCAGT TGTGCAGGGG GGACCAGGGT GGGGAAAGTT CACTGTCACT | 1800 |
| | CTCCAAGTGG AACACTTTTC TGAAAGCCAT GCTGGTATGC AGTGATGCTG CCACCAACAA | 1860 |
| 15 | GAACTTCAAC AGGCTGCAAG ACGTCTTCCT GCTCCCTGAC CCCAGCGGCC AGTGGAGGGA | 1920 |
| | CACCAGGGTC TATGGTGTTC TCTCCAACCC CTGGAAGTAC TCAGCCGTCT GTGTGTATTC | 1980 |
| | CCTCGGTGAC ATTGACAAGG TCTTCCGTAC CTCCTCACTC AAGGGCTACC ACTCAAGCCT | 2040 |
| 20 | TCCCAACCCG CGGCCTGGCA AGTGCCTCCC AGACCAGCAG CCGATACCCA CAGAGACCTT | 2100 |
| | CCAGGTGGCT GACCGTCACC CAGAGGTGGC GCAGAGGGTG GAGCCCATGG GGCCTCTGAA | 2160 |
| 25 | GACGCCATTG TTCCACTCTA AATACCACTA CCAGAAAGTG GCCGTTACC GCATGCAAGC | 2220 |
| | CAGCCACGGG GAGACCTTTC ATGTGCTTTA CCTAACTACA GACAGGGGCA CTATCCACAA | 2280 |
| | GGTGGTGGAA CCGGGGGAGC AGGAGCACAG CTTGCGCTTC AACATCATGG AGATCCAGCC | 2340 |
| 30 | CTTCCGCCGC GCGGCTGCCA TCCAGACCAT GTCGCTGGAT GCTGAGCGGA GGAAGCTGTA | 2400 |
| | TGTGAGCTCC CAGTGGGAGG TGAGCCAGGT GCCCTGGAC CTGTGTGAGG TCTATGGCGG | 2460 |
| | GGGCTGCCAC GGTTGCCTCA TGTCCCGAGA CCCCTACTGC GGCTGGGACC AGGGCCGCTG | 2520 |
| 35 | CATCTCCATC TACAGCTCCG AACGGTCAGT GCTGCAATCC ATTAATCCAG CCGAGCCACA | 2580 |
| | CAAGGAGTGT CCCAACCCCA AACCAGACAA GGCCCCACTG CAGAAGGTTT CCCTGGCCCC | 2640 |
| 40 | AAACTCTCGC TACTACCTGA GCTGCCCCAT GGAATCCCGC CACGCCACCT ACTCATGGCG | 2700 |
| | CCACAAGGAG AACGTGGAGC AGAGCTGCGA ACCTGGTCAC CAGAGCCCCA ACTGCATCCT | 2760 |
| | GTTTCATCGAG AACCTCACGG CGCAGCAGTA CGGCCACTAC TTCTGCGAGG CCCAGGAGGG | 2820 |
| 45 | CTCCTACTTC CGCGAGGCTC AGCACTGGCA GCTGCTGCCC GAGGACGGCA TCATGGCCGA | 2880 |
| | GCACCTGCTG GGTTCATGCCT GTGCCCTGGC TGCCTCCCTC TGGCTGGGGG TGCTGCCAC | 2940 |
| | ACTCACTCTT GGCTTGCTGG TCCACATGGT GAGCAAGGGC GAGGAGCTGT TCACCGGGT | 3000 |
| 50 | GGTGCCCATC CTGGTCGAGC TGGACGGCGA CGTAAACGGC CACAAGTTCA GCGTGTCCGG | 3060 |
| | CGAGGGCGAG GCGATGCCA CCTACGGCAA GCTGACCCTG AAGTTCATCT GCACCACCGG | 3120 |

| | | |
|----|---|------|
| | CAAGCTGCCC GTGCCCTGGC CCACCCTCGT GACCACCCTG ACCTACGGCG TGCAGTGCTT | 3180 |
| 5 | CAGCCGCTAC CCCGACCACA TGAAGCAGCA CGACTTCTTC AAGTCCGCCA TGCCCGAAGG | 3240 |
| | CTACGTCCAG GAGCGCACCA TCTTCTTCAA GGACGACGGC AACTACAAGA CCCGCGCCGA | 3300 |
| | GGTGAAGTTC GAGGGCGACA CCCTGGTGAA CCGCATCGAG CTGAAGGGCA TCGACTTCAA | 3360 |
| 10 | GGAGGACGGC AACATCCTGG GGCACAAGCT GGAGTACAAC TACAACAGCC ACAACGTCTA | 3420 |
| | TATCATGGCC GACAAGCAGA AGAACGGCAT CAAGGTGAAC TTCAAGATCC GCCACAACAT | 3480 |
| | CGAGGACGGC AGCGTGCAGC TCGCCGACCA CTACCAGCAG AACACCCCCA TCGGCGACGG | 3540 |
| 15 | CCCCGTGCTG CTGCCCAGCA ACCACTACCT GAGCACCCAG TCCGCCCTGA GCAAAGACCC | 3600 |
| | CAACGAGAAG CGCGATCACA TGGTCCTGCT GGAGTTCGTG ACCGCCGCCG GGATCACTCT | 3660 |
| 20 | CGGCATGGAC GAGCTGTACA AGGTGAAGCT TGGGCCCCGAA CAAAACTCA TCTCAGAAGA | 3720 |
| | GGATCTGAAT AGCGCCGTGG ACCATCATCA TCATCATCAT TGAGTTTAAA CCGCTGATCA | 3780 |
| | GCCTCGACTG TGCCTTCTAG TTGCCAGCCA TCTGTTGTTT GCCCCTCCCC CGTGCCTTCC | 3840 |
| 25 | TTGACCCTGG AAGGTGCCAC TCCCACTGTC CTTTCCTAAT AAAATGAGGA AATTGCATCG | 3900 |
| | CATTGTCTGA GTAGGTGTCA TTCTATTCTG GGGGGTGGGG TGGGGCAGGA CAGCAAGGGG | 3960 |
| 30 | GAGGATTGGG AAGACAATAG CAGGCATGCT GGGGATGCGG TGGGCTCTAT GGCTTCTGAG | 4020 |
| | GCGGAAAGAA CCAGCTGGGG CTCTAGGGGG TATCCCCACG CGCCCTGTAG CGGCGCATT | 4080 |
| | AGCGCGGCGG GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG | 4140 |
| 35 | CCCGCTCCTT TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA | 4200 |
| | GCTCTAAATC GGGGCATCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC | 4260 |
| | AAAAAACTTG ATTAGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT | 4320 |
| 40 | CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGGAACA | 4380 |
| | AACTCAACC CTATCTCGGT CTATTCTTTT GATTTATAAG GGATTTTGGG GATTTCGGCC | 4440 |
| 45 | TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAAGC CGAATTAATT CTGTGGAATG | 4500 |
| | TGTGTCAGTT AGGGTGTGGA AAGTCCCCAG GCTCCCCAGG CAGGCAGAAG TATGCAAAGC | 4560 |
| | ATGCATCTCA ATTAGTCAGC AACCAGGTGT GGAAAGTCCC CAGGCTCCCC AGCAGGCAGA | 4620 |
| 50 | AGTATGCAAA GCATGCATCT CAATTAGTCA GCAACCATAG TCCCGCCCCT AACTCCGCCC | 4680 |
| | ATCCCGCCCC TAACTCCGCC CAGTTCGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT | 4740 |

TTTATTTATG CAGAGGCCGA GGCCGCCTCT GCCTCTGAGC TATTCCAGAA GTAGTGAGGA 4800
 GGCTTTTTTG GAGGCCTAGG CTTTGTCAAA AAGCTCCCGG GAGCTTGTAT ATCCATTTTC 4860
 5 GGATCTGATC AAGAGACAGG ATGAGGATCG TTTCGCATGA TTGAACAAGA TGGATTGCAC 4920
 GCAGGTTCTC CGGCCGCTTG GGTGGAGAGG CTATTCCGCT ATGACTGGGC ACAACAGACA 4980
 10 ATCGGCTGCT CTGATGCCGC CGTGTTCCGG CTGTCAGCGC AGGGGCGCCC GGTTCTTTTT 5040
 GTCAAGACCG ACCTGTCCGG TGCCCTGAAT GAACTGCAGG ACGAGGCAGC GCGGCTATCG 5100
 TGGCTGGCCA CGACGGGCGT TCCTTGCGCA GCTGTGCTCG ACGTTGTAC TGAAGCGGGA 5160
 15 AGGGA CTGGC TGCTATTGGG CGAAGTGCCG GGGCAGGATC TCCTGTATC TCACCTTGCT 5220
 CCTGCCGAGA AAGTATCCAT CATGGCTGAT GCAATGCCGC GGCTGCATAC GCTTGATCCG 5280
 GCTACCTGCC CATTGACCA CCAAGCGAAA CATCGCATCG AGCGAGCAGC TACTCGGATG 5340
 20 GAAGCCGGTC TTGTCGATCA GGATGATCTG GACGAAGAGC ATCAGGGGCT CGCGCCAGCC 5400
 GAACTGTTCTG CCAGGCTCAA GGCGCGCATG CCCGACGGCG AGGATCTCGT CGTGACCCAT 5460
 GGCGATGCCT GCTTGCCGAA TATCATGGTG GAAAATGGCC GCTTTTCTGG ATTCATCGAC 5520
 25 TGTGGCCGGC TGGGTGTGGC GGACCGCTAT CAGGACATAG CGTTGGCTAC CCGTGATATT 5580
 GCTGAAGAGC TTGGCGGCGA ATGGGCTGAC CGCTTCCTCG TGCTTTACGG TATCGCCGCT 5640
 30 CCCGATTGCG AGCGCATCGC CTTCTATCGC CTTCTTGACG AGTTCTTCTG AGCGGGACTC 5700
 TGGGGTTCGA AATGACCGAC CAAGCGACGC CCAACCTGCC ATCACCAGAT TTCGATTCCA 5760
 CCGCCGCCCTT CTATGAAAGG TTGGGCTTCG GAATCGTTTT CCGGGACGCC GGCTGGATGA 5820
 35 TCCTCCAGCG CGGGGATCTC ATGCTGGAGT TCTTCGCCCA CCCCAACTTG TTTATTGCAG 5880
 CTTATAATGG TTACAAATAA AGCAATAGCA TCACAAATTT CACAAATAAA GCATTTTTTT 5940
 CACTGCATTG TAGTTGTGGT TTGTCCAAAC TCATCAATGT ATCTTATCAT GTCTGTATAC 6000
 40 CGTCGACCTC TAGCTAGAGC TTGGCGTAAT CATGGTCATA GCTGTTTCCT GTGTGAAATT 6060
 GTTATCCGCT CACAATTCCA CACAACATAC GAGCCGGAAG CATAAAGTGT AAAGCCTGGG 6120
 GTGCCTAATG AGTGAGCTAA CTCACATTAA TTGCGTTGCG CTCACTGCCC GCTTTCCAGT 6180
 45 CGGGAAACCT GTCGTGCCAG CTGCATTAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT 6240
 TGCGTATTGG GCGCTCTTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG GTCGTTCCGC 6300
 50 TGCGGCGAGC GGTATCAGCT CACTCAAAGG CGGTAATACG GTTATCCACA GAATCAGGGG 6360
 ATAACGCAGG AAAGAACATG TGAGCAAAAG GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG 6420

CCGCGTTGCT GCGGTTTTTC CATAGGCTCC GCCCCCTGA CGAGCATCAC AAAAATCGAC 6480
 GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG TTCCCCCTG 6540
 5 GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA CCCTGCCGCT TACCGGATAC CTGTCCGCCT 6600
 TTCTCCCTTC GGAAGCGTG GCGCTTTCTC AATGCTCAG CTGTAGGTAT CTCAGTTCGG 6660
 10 TGTAGGTCGT TCGCTCCAAG CTGGGCTGTG TGCACGAACC CCCCCTTCAG CCCGACCGCT 6720
 GCGCCTTATC CGGTAACAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC 6780
 TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT GCTACAGAGT 6840
 15 TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGGAC AGTATTTGGT ATCTGCGCTC 6900
 TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC TTGATCCGGC AAACAAACCA 6960
 CCGCTGGTAG CGGTGGTTTT TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT 7020
 20 CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGAAC GAAAACTCAC 7080
 GTTAAGGGAT TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAATT 7140
 25 AAAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA AACTTGGTCT GACAGTTACC 7200
 AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTCTGTCA TCCATAGTTG 7260
 CCTGACTCCC CGTCGTGTAG ATAACCTACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG 7320
 30 CTGCAATGAT ACCGCGAGAC CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC 7380
 CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT ATCCGCCTCC ATCCAGTCTA 7440
 35 TTAATTGTTG CCGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG 7500
 TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTCGTT TGGTATGGCT TCATTAGCT 7560
 CCGGTTCCCA ACGATCAAGG CGAGTTACAT GATCCCCCAT GTTGTGCAAA AAAGCGGTTA 7620
 40 GCTCCTTCGG TCCTCCGATC GTTGTGAGAA GTAAGTTGGC CGCAGTGTTA TCACTCATGG 7680
 TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA 7740
 CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGCGACCG AGTTGCTCTT 7800
 45 GCCCGGCGTC AATACGGGAT AATACCGCGC CACATAGCAG AACTTTAAAA GTGCTCATCA 7860
 TTGGAAAACG TTCTTCGGGG CGAAAACTCT CAAGGATCTT ACCGCTGTTG AGATCCAGTT 7920
 50 CGATGTAACC CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT 7980
 CTGGGTGAGC AAAACAGGA AGGCAAAATG CCGCAAAAAA GGAATAAGG GCGACACGGA 8040

AATGTTGAAT ACTCATACTC TTCCTTTTTC AATATTATTG AAGCATTTAT CAGGGTTATT 8100
 GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA TAAACAAATA GGGGTTCCGC 8160
 GCACATTTCC CCGAAAAGTG CCACCTGACG TC 8192

(2) INFORMATION ZU SEQ ID NO: 37:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 7000 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LÄNGE: 1..7000

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 37:

AGATCTCGGC CGCATATTAA GTGCATTGTT CTCGATACCG CTAAGTGCAT TGTTCTCGTT 60
 AGCTCGATGG ACAAGTGCAT TGTTCTCTTG CTGAAAGCTC GATGGACAAG TGCATTGTTC 120
 TCTTGCTGAA AGCTCGATGG ACAAGTGCAT TGTTCTCTTG CTGAAAGCTC AGTACCCGGG 180
 AGTACCCTCG ACCGCCGGAG TATAAATAGA GGCGCTTCGT CTACGGAGCG ACAATTCAAT 240
 TCAAACAAGC AAAGTGAACA CGTCGCTAAG CGAAAGCTAA GCAAATAAAC AAGCGCAGCT 300
 GAACAAGCTA AACAATCTGC AGTAAAGTGC AAGTTAAAGT GAATCAATTA AAAGTAACCA 360
 GCAACCAAGT AAATCAACTG CAACTACTGA AATCTGCCAA GAAGTAATTA TTGAATACAA 420
 GAAGAGAACT CTGAATACTT TCAACAAGTT ACCGAGAAAG AAGAACTCAC ACACAGCTAG 480
 CGTTTAAACT TAAGCTTGGT ACCGAGCTCG GATCCACTAG TCCAGTGTGG TGGAAATTCGG 540
 CTTGGGATGA CGCCTCCTCC GCCCGGACGT GCCGCCCCCA GCGCACC GCG CCCCCGCTC 600
 CCTGGCCCCG CGGCTCGGTT GGGGCTTCCG CTGCGGCTGC GGCTGCTGCT GCTGCTCTGG 660
 GCGGCCGCCG CCTCCGCCCA GGGCCACCTA AGGAGCGGAC CCCGCATCTT CGCCGTCTGG 720
 AAAGGCCATG TAGGGCAGGA CCGGGTGGAC TTTGGCCAGA CTGAGCCGCA CACGGTGCTT 780
 TTCCACGAGC CAGGCAGCTC CTCTGTGTGG GTGGGAGGAC GTGGCAAGGT CTACCTCTTT 840
 GACTTCCCCG AGGGCAAGAA CGCATCTGTG CGCACGGTGA ATATCGGCTC CACAAAGGGG 900

TCCTGTCTGG ATAAGCGGGA CTGCGAGAAC TACATCACTC TCCTGGAGAG GCGGAGTGAG 960
 GGGCTGCTGG CCTGTGGCAC CAACGCCCCG CACCCAGCT GCTGGAACCT GGTGAATGGC 1020
 5 ACTGTGGTGC CACTTGGCGA GATGAGAGGC TACGCCCCCT TCAGCCCGGA CGAGAACTCC 1080
 CTGGTTCTGT TTGAAGGGGA CGAGGTGTAT TCCACCATCC GGAAGCAGGA ATACAATGGG 1140
 AAGATCCCTC GGTTCCGCCG CATCCGGGGC GAGAGTGAGC TGTACACCAG TGATACTGTC 1200
 10 ATGCAGAACC CACAGTTCAT CAAAGCCACC ATCGTGCACC AAGACCAGGC TTACGATGAC 1260
 AAGATCTACT ACTTCTTCCG AGAGGACAAT CCTGACAAGA ATCCTGAGGC TCCTCTCAAT 1320
 15 GTGTCCCGTG TGGCCAGTT GTGCAGGGGG GACCAGGGTG GGGAAAGTTC ACTGTCAGTC 1380
 TCCAAGTGGG AACCTTTTCT GAAAGCCATG CTGGTATGCA GTGATGCTGC CACCAACAAG 1440
 AACTTCAACA GGCTGCAAGA CGTCTTCCTG CTCCCTGACC CCAGCGGCCA GTGGAGGGAC 1500
 20 ACCAGGGTCT ATGGTGTITT CTCCAACCCC TGGAACTACT CAGCCGTCTG TGTGTATTCC 1560
 CTCGGTGACA TTGACAAGGT CTTCCGTACC TCCTCACTCA AGGGCTACCA CTCAAGCCTT 1620
 CCCAACCCGC GGCCTGGCAA GTGCCTCCCA GACCAGCAGC CGATACCCAC AGAGACCTTC 1680
 25 CAGGTGGCTG ACCGTCACCC AGAGGTGGCG CAGAGGGTGG AGCCCATGGG GCCTCTGAAG 1740
 ACGCCATTGT TCCACTCTAA ATACCACTAC CAGAAAGTGG CCGTTCACCG CATGCAAGCC 1800
 30 AGCCACGGGG AGACCTTTCA TGTGCTTTAC CTAACCTACAG ACAGGGGCAC TATCCACAAG 1860
 GTGGTGGAAC CGGGGAGCA GGAGCACAGC TTCGCCTTCA ACATCATGGA GATCCAGCCC 1920
 TTCCGCCGCG CGGCTGCCAT CCAGACCATG TCGCTGGATG CTGAGCGGAG GAAGCTGTAT 1980
 35 GTGAGCTCCC AGTGGGAGGT GAGCCAGGTG CCCCTGGACC TGTGTGAGGT CTATGGCGGG 2040
 GGCTGCCACG GTTGCCTCAT GTCCCGAGAC CCCTACTGCG GCTGGGACCA GGGCCGCTGC 2100
 ATCTCCATCT ACAGCTCCGA ACGGTCAGTG CTGCAATCCA TTAATCCAGC CGAGCCACAC 2160
 40 AAGGAGTGTC CCAACCCCAA ACCAGACAAG GCCCCTGTC AGAAGGTTTC CCTGGCCCCA 2220
 AACTCTCGCT ACTACCTGAG CTGCCCCATG GAATCCCGCC ACGCCACCTA CTCATGGCGC 2280
 45 CACAAGGAGA ACGTGGAGCA GAGCTGCGAA CCTGGTCACC AGAGCCCCAA CTGCATCCTG 2340
 TTCATCGAGA ACCTCACGGC GCAGCAGTAC GGCCACTACT TCTGCGAGGC CCAGGAGGGC 2400
 TCCTACTTCC GCGAGGCTCA GCACTGGCAG CTGCTGCCCG AGGACGGCAT CATGGCCGAG 2460
 50 CACCTGCTGG GTCATGCCTG TGCCCTGGCT GCCTCCCTCT GGCTGGGGGT GCTGCCACCA 2520
 CTCACTCTTG GCTTGCTGGT CCACGTGAAG CTGGGGCCCG TTAAACCCG CTGATCAGCC 2580

| | | |
|----|--|------|
| | TCGACTGTGC CTTCTAGTTG CCAGCCATCT GTTGTGTTGCC CCTCCCCCGT GCCTTCCTTG | 2640 |
| | ACCCTGGAAG GTGCCACTCC CACTGTCCTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT | 2700 |
| 5 | TGTCTGAGTA GGTGTCATTG TATTCTGGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG | 2760 |
| | GATTGGGAAG ACAATAGCAG GCATGCTGGG GATGCCGTGG GCTCTATGGC TTCTGAGGCG | 2820 |
| 10 | GAAAGAACCA GCTGGGGCTC TAGGGGGTAT CCCCACGCGC CCTGTAGCGG CGCATTAAAGC | 2880 |
| | GCGGCGGGTG TGGTGGTTAC GCGCAGCGTG ACCGCTACAC TTGCCAGCGC CCTAGCGCCC | 2940 |
| | GCTCCTTTTCG CTTTCTTCCC TTCCTTTCTC GCCACGTTTC CCGGCTTTCC CCGTCAAGCT | 3000 |
| 15 | CTAAATCGGG GCATCCCTTT AGGGTTCCGA TTTAGTGCTT TACGGCACCT CGACCCCAAA | 3060 |
| | AAACTTGATT AGGGTGATGG TTCACGTAGT GGGCCATCGC CCTGATAGAC GGTTTTTTCG | 3120 |
| 20 | CCTTTGACGT TGGAGTCCAC GTTCTTTAAT AGTGGACTCT TGTTCCAAAC TGGAACAACA | 3180 |
| | CTCAACCCTA TCTCGGTCTA TTCTTTTGAT TTATAAGGGA TTTTGGGGAT TTCGGCCTAT | 3240 |
| | TGGTTAAAAA ATGAGCTGAT TTAACAAAAA TTTAACGCGA ATTAATTCTG TGGAATGTGT | 3300 |
| 25 | GTCAGTTAGG GTGTGGAAAG TCCCCAGGCT CCCCAGGCAG GCAGAAGTAT GCAAAGCATG | 3360 |
| | CATCTCAATT AGTCAGCAAC CAGGTGTGGA AAGTCCCCAG GCTCCCCAGC AGGCAGAAAT | 3420 |
| | ATGCAAAGCA TGCATCTCAA TTAGTCAGCA ACCATAGTCC CGCCCCTAAC TCCGCCCATC | 3480 |
| 30 | CCGCCCCCTAA CTCCGCCAG TTCCGCCCAT TCTCCGCCCC ATGGCTGACT AATTTTTTTT | 3540 |
| | ATTTATGCAG AGGCCGAGGC CGCCTCTGCC TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC | 3600 |
| 35 | TTTTTTGGAG GCCTAGGCTT TTGCAAAAAG CTCCCGGGAG CTTGTATATC CATTTTCGGA | 3660 |
| | TCTGATCAAG AGACAGGATG AGGATCGTTT CGCATGATTG AACAAGATGG ATTGCACGCA | 3720 |
| | GGTTCTCCGG CCGCTTGGGT GGAGAGGCTA TTCGGCTATG ACTGGGCACA ACAGACAATC | 3780 |
| 40 | GGCTGCTCTG ATGCCGCCGT GTTCCGGCTG TCAGCGCAGG GGCGCCCGGT TCTTTTTGTC | 3840 |
| | AAGACCGACC TGTCCGGTGC CCTGAATGAA CTGCAGGACG AGGCAGCGCG GCTATCGTGG | 3900 |
| 45 | CTGGCCACGA CGGGCGTTCC TTGCGCAGCT GTGCTCGACG TTGTCACTGA AGCGGGAAGG | 3960 |
| | GACTGGCTGC TATTGGGCGA AGTGCCGGGG CAGGATCTCC TGTCATCTCA CCTTGCTCCT | 4020 |
| | GCCGAGAAAG TATCCATCAT GGCTGATGCA ATGCGGCGGC TGCATACGCT TGATCCGGCT | 4080 |
| 50 | ACCTGCCCCAT TCGACCACCA AGCGAAACAT CGCATCGAGC GAGCACGTAC TCGGATGGAA | 4140 |
| | GCCGGTCTTG TCGATCAGGA TGATCTGGAC GAAGAGCATC AGGGGCTCGC GCCAGCCGAA | 4200 |

CTGTTGCGCA GGCTCAAGGC GCGCATGCCC GACGGCGAGG ATCTCGTCGT GACCCATGGC 4260
 GATGCCTGCT TGCCGAATAT CATGGTGGAA AATGGCCGCT TTTCTGGATT CATCGACTGT 4320
 5 GGCCGGCTGG GTGTGGCGGA CCGCTATCAG GACATAGCGT TGGCTACCCG TGATATTGCT 4380
 GAAGAGCTTG GCGGCGAATG GGCTGACCGC TTCCTCGTGC TTTACGGTAT CGCCGCTCCC 4440
 10 GATTGCGAGC GCATCGCCTT CTATCGCCTT CTTGACGAGT TCTTCTGAGC GGGACTCTGG 4500
 GGTTCGAAAT GACCGACCAA GCGACGCCCC ACCTGCCATC ACGAGATTTC GATTCCACCG 4560
 CCGCCTTCTA TGAAAGGTTG GGCTTCGGAA TCGTTTTCCG GGACGCCGGC TGGATGATCC 4620
 15 TCCAGCGCGG GGATCTCATG CTGGAGTTCT TCGCCACCC CAACTTGTTT ATTGCAGCTT 4680
 ATAATGGTTA CAAATAAAGC AATAGCATCA CAAATTTTAC AAATAAAGCA TTTTTTTTAC 4740
 TGCATTCTAG TTGTGGTTTG TCCAACTCA TCAATGTATC TTATCATGTC TGTATACCGT 4800
 20 CGACCTCTAG CTAGAGCTTG GCGTAATCAT GGTTCATAGT GTTTCCTGTG TGAAATTGTT 4860
 ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAG GCCTGGGGTG 4920
 CCTAATGAGT GAGCTAACTC ACATTAATTG CGTTGCGCTC ACTGCCCCGT TTCCAGTCGG 4980
 25 GAAACCTGTC GTGCCAGCTG CATTAAATGAA TCGGCCAACG CGCGGGGAGA GGCGGTTTGC 5040
 GTATTGGGCG CTCTTCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC 5100
 30 GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA TCAGGGGATA 5160
 ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC CAGGAACCGT AAAAAGGCCG 5220
 CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC CCCCTGACGA GCATCACAAA AATCGACGCT 5280
 35 CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCCTGGAA 5340
 GCTCCCTCGT GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCCTTTC 5400
 TCCCTTCGGG AAGCGTGGCG CTTTCTCAAT GCTCACGCTG TAGGTATCTC AGTTCGGTGT 5460
 40 AGGTCGTTTC CTCCAAGCTG GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCGCTGCG 5520
 CCTTATCCGG TAACTATCGT CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG 5580
 CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT 5640
 45 TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TGCGCTCTGC 5700
 TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA CAAACCACCG 5760
 50 CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC 5820
 AAGAAGATCC TTTGATCTTT TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCACGTT 5880

AAGGGATTTT GGTCATGAGA TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA 5940
 AATGAAGTTT TAAATCAATC TAAAGTATAT ATGACTAAAC TTGGTCTGAC AGTTACCAAT 6000
 5 GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC ATAGTTGCCT 6060
 GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT ACCATCTGGC CCCAGTGTCTG 6120
 10 CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT ATCAGCAATA AACCAGCCAG 6180
 CCGGAAGGGC CGAGCGCAGA AGTGGTCTCTG CAACTTTATC CGCCTCCATC CAGTCTATTA 6240
 ATTGTTGCCG GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTTGCGC AACGTTGTTG 6300
 15 CCATTGCTAC AGGCATCGTG GTGTCACGCT CGTCGTTTGG TATGGCTTCA TTCAGCTCCG 6360
 GTTCCCAACG ATCAAGGCGA GTTACATGAT CCCCCATGTT GTGCAAAAAA GCGGTTAGCT 6420
 20 CCTTCGGTCC TCCGATCGTT GTCAGAAGTA AGTTGGCCGC AGTGTTATCA CTCATGGTTA 6480
 TGGCAGCACT GCATAATTCT CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG 6540
 GTGAGTACTC AACCAAGTCA TTCTGAGAAT AGTGATGCG GCGACCGAGT TGCTCTTGCC 6600
 25 CGGCGTCAAT ACGGGATAAT ACCGCGCCAC ATAGCAGAAC TTAAAAGTG CTCATCATTG 6660
 GAAAACGTTT TTCGGGGCGA AACTCTCAA GGATCTTACC GCTGTTGAGA TCCAGTTCGA 6720
 30 TGTAACCCAC TCGTGCACCC AACTGATCTT CAGCATCTTT TACTTTCACC AGCGTTTCTG 6780
 GGTGAGCAAA AACAGGAAGG CAAAATGCCG CAAAAAGGG AATAAGGGCG ACACGGAAAT 6840
 GTTGAATACT CATACTCTTC CTTTTTCAAT ATTATTGAAG CATTTATCAG GGTTATTGTC 6900
 35 TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA ACAAATAGGG GTTCCGCGCA 6960
 CATTTCCCCG AAAAGTGCCA CCTGACGTCG ACGGATCGGG 7000

40 (2) INFORMATION ZU SEQ ID NO: 38:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 7108 Basenpaare
 (B) ART: Nukleinsäure
 45 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

50 (ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
 (B) LÄGE: 1..7108

55

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 38:

| | | |
|----|--|------|
| 5 | AGATCTCGGC CGCATATTAA GTGCATTGTT CTCGATACCG CTAAGTGCAT TGTTCCTCGTT | 60 |
| | AGCTCGATGG ACAAGTGCAT TGTTCCTTTG CTGAAAGCTC GATGGACAAG TGCATTGTTC | 120 |
| | TCTTGCTGAA AGCTCGATGG ACAAGTGCAT TGTTCCTTTG CTGAAAGCTC AGTACCCGGG | 180 |
| 10 | AGTACCCTCG ACCGCCGGAG TATAAATAGA GGCCTTCGT CTACGGAGCG ACAATTCAAT | 240 |
| | TCAAACAAGC AAAGTGAACA CGTCGCTAAG CGAAAGCTAA GCAAATAAAC AAGCGCAGCT | 300 |
| | GAACAAGCTA AACAATCTGC AGTAAAGTGC AAGTTAAAGT GAATCAATTA AAAGTAACCA | 360 |
| 15 | GCAACCAAGT AAATCAACTG CAACTACTGA AATCTGCCAA GAAGTAATTA TTGAATACAA | 420 |
| | GAAGAGAACT CTGAATACTT TCAACAAGTT ACCGAGAAAG AAGAACTCAC ACACAGCTAG | 480 |
| 20 | CGTTTAAACT TAAGCTTGGT ACCGAGCTCG GATCCACTAG TCCAGTGTGG TGAATTCCGG | 540 |
| | CTTGGGATGA CGCCTCCTCC GCCCGGACGT GCCGCCCCCA GCGCACCGCG CGCCCGCGTC | 600 |
| | CCTGGCCCCG CGGCTCGGTT GGGGCTTCCG CTGCGGCTGC GGCTGCTGCT GCTGCTCTGG | 660 |
| 25 | GCGGCCGCCG CCTCCGCCCA GGGCCACCTA AGGAGCGGAC CCCGCATCTT CGCCGTCTGG | 720 |
| | AAAGGCCATG TAGGGCAGGA CCGGGTGGAC TTTGGCCAGA CTGAGCCGCA CACGGTGCTT | 780 |
| 30 | TTCCACGAGC CAGGCAGCTC CTCTGTGTGG GTGGGAGGAC GTGGCAAGGT CTACCTCTTT | 840 |
| | GACTTCCCCG AGGGCAAGAA CGCATCTGTG CGCACGGTGA ATATCGGCTC CACAAAGGGG | 900 |
| | TCCTGTCTGG ATAAGCGGGA CTGCGAGAAC TACATCACTC TCCTGGAGAG GCGGAGTGAG | 960 |
| 35 | GGGCTGCTGG CCTGTGGCAC CAACGCCCCG CACCCAGCT GCTGGAACCT GGTGAATGGC | 1020 |
| | ACTGTGGTGC CACTTGGCGA GATGAGAGGC TACGCCCCCT TCAGCCCGGA CGAGAACTCC | 1080 |
| | CTGGTTCTGT TTGAAGGGGA CGAGGTGTAT TCCACCATCC GGAAGCAGGA ATACAATGGG | 1140 |
| 40 | AAGATCCCTC GGTTCGCCCG CATCCGGGGC GAGAGTGAGC TGTACACCAG TGATACTGTC | 1200 |
| | ATGCAGAACC CACAGTTCAT CAAAGCCACC ATCGTGACC AAGACCAGGC TTACGATGAC | 1260 |
| 45 | AAGATCTACT ACTTCTTCCG AGAGGACAAT CCTGACAAGA ATCCTGAGGC TCCTCTCAAT | 1320 |
| | GTGTCCCGTG TGGCCAGTT GTGCAGGGGG GACCAGGGTG GGGAAAGTTC ACTGTCAGTC | 1380 |
| | TCCAAGTGA ACACTTTTCT GAAAGCCATG CTGGTATGCA GTGATGCTGC CACCAACAAG | 1440 |
| 50 | AACTTCAACA GGCTGCAAGA CGTCTTCCTG CTCCCTGACC CCAGCGGCCA GTGGAGGGAC | 1500 |
| | ACCAGGGTCT ATGGTGT TTT CTCCAACCCC TGGAACTACT CAGCCGTCTG TGTGTATTCC | 1560 |

| | | |
|----|---|------|
| | CTCGGTGACA TTGACAAGGT CTTCCGTACC TCCTCACTCA AGGGCTACCA CTCAAGCCTT | 1620 |
| 5 | CCCAACCCGC GGCCTGGCAA GTGCCTCCCA GACCAGCAGC CGATACCCAC AGAGACCTTC | 1680 |
| | CAGGTGGCTG ACCGTCACCC AGAGGTGGCG CAGAGGGTGG AGCCCATGGG GCCTCTGAAG | 1740 |
| | ACGCCATTGT TCCACTCTAA ATACCACTAC CAGAAAGTGG CCGTTCACCG CATGCAAGCC | 1800 |
| 10 | AGCCACGGGG AGACCTTTCA TGTGCTTTAC CTAATACTACAG ACAGGGGCAC TATCCACAAG | 1860 |
| | GTGGTGGAAAC CGGGGGAGCA GGAGCACAGC TTCGCCTTCA ACATCATGGA GATCCAGCCC | 1920 |
| | TTCCGCCGCG CGGCTGCCAT CCAGACCATG TCGCTGGATG CTGAGCGGAG GAAGCTGTAT | 1980 |
| 15 | GTGAGCTCCC AGTGGGAGGT GAGCCAGGTG CCCCTGGACC TGTGTGAGGT CTATGGCGGG | 2040 |
| | GGCTGCCACG GTTGCTCAT GTCCCGAGAC CCCTACTGCG GCTGGGACCA GGGCCGCTGC | 2100 |
| | ATCTCCATCT ACAGCTCCGA ACGGTCAGTG CTGCAATCCA TTAATCCAGC CGAGCCACAC | 2160 |
| 20 | AAGGAGTGTC CCAACCCCAA ACCAGACAAG GCCCCACTGC AGAAGGTTTC CCTGGCCCCA | 2220 |
| | AACTCTCGCT ACTACCTGAG CTGCCCCATG GAATCCCGCC ACGCCACCTA CTCATGGCGC | 2280 |
| 25 | CACAAGGAGA ACGTGGAGCA GAGCTGCGAA CCTGGTCACC AGAGCCCCAA CTGCATCCTG | 2340 |
| | TTCATCGAGA ACCTCACGGC GCAGCAGTAC GGCCACTACT TCTGCGAGGC CCAGGAGGGC | 2400 |
| | TCCTACTTCC GCGAGGCTCA GCACTGGCAG CTGCTGCCCC AGGACGGCAT CATGGCCGAG | 2460 |
| 30 | CACCTGCTGG GTCATGCCTG TGCCCTGGCT GCCTCCCTCT GGCTGGGGGT GCTGCCCA | 2520 |
| | CTCACTCTTG GCTTGCTGGT CCACGTGAAG CTTGGGCCCC AACAAAACT CATCTCAGAA | 2580 |
| | GAGGATCTGA ATAGCGCCGT CGACCATCAT CATCATCATC ATTGAGTTTA TCCAGCACAG | 2640 |
| 35 | TGGCGGCCGC TCGAGTCTAG AGGGCCCGTT TAAACCCGCT GATCAGCCTC GACTGTGCCT | 2700 |
| | TCTAGTTGCC AGCCATCTGT TGTTTGCCCC TCCCCCGTGC CTTCTTGAC CCTGGAAGGT | 2760 |
| 40 | GCCACTCCCA CTGTCTTTTC CTAATAAAAT GAGGAAATTG CATCGCATTG TCTGAGTAGG | 2820 |
| | TGTCATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA AGGGGGAGGA TTGGGAAGAC | 2880 |
| | AATAGCAGGC ATGCTGGGGA TGCGGTGGGC TCTATGGCTT CTGAGGCGGA AAGAACCAGC | 2940 |
| 45 | TGGGGCTCTA GGGGGTATCC CCACGCGCCC TGTAGCGGCG CATTAAAGCGC GGCGGGTGTG | 3000 |
| | GTGGTTACGC GCAGCGTGAC CGCTACACTT GCCAGCGCCC TAGCGCCCGC TCCTTTCGCT | 3060 |
| | TTCTTCCCTT CCTTCTCGC CACGTTCCGC GGCTTTCCTT GTCAAGCTCT AAATCGGGG | 3120 |
| 50 | ATCCCTTTAG GGTTCGATT TAGTGCTTTA CGGCACCTCG ACCCCAAAAA ACTTGATTAG | 3180 |
| | GGTGATGGTT CACGTAGTGG GCCATCGCCC TGATAGACGG TTTTTCGCCC TTTGACGTTG | 3240 |

GAGTCCACGT TCTTTAATAG TGGACTCTTG TTCCAAACTG GAACAACACT CAACCCTATC 3300
 TCGGTCTATT CTTTTGATTT ATAAGGGATT TTGGGGATTG CGGCCTATTG GTTAAAAAAT 3360
 5 GAGCTGATTT AACAAAAATT TAACGCGAAT TAATTCTGTG GAATGTGTGT CAGTTAGGGT 3420
 GTGGAAAGTC CCCAGGCTCC CCAGGCAGGC AGAAGTATGC AAAGCATGCA TCTCAATTAG 3480
 10 TCAGCAACCA GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG 3540
 CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC CGCCCATCCC GCCCCTAACT 3600
 CCGCCAGTT CCGCCATTG TCCGCCCCAT GGCTGACTAA TTTTTTTTAT TTATGCAGAG 3660
 15 GCCGAGGCCG CCTCTGCCTC TGAGCTATTC CAGAAGTAGT GAGGAGGCTT TTTTGGAGGC 3720
 CTAGGCTTTT GCAAAAAGCT CCCGGGAGCT TGTATATCCA TTTTCGGATC TGATCAAGAG 3780
 ACAGGATGAG GATCGTTTCG CATGATTGAA CAAGATGGAT TGCACGCAGG TTCTCCGGCC 3840
 20 GCTTGGGTGG AGAGGCTATT CGGCTATGAC TGGGCACAAC AGACAATCGG CTGCTCTGAT 3900
 GCCGCCGTGT TCCGGCTGTC AGCGCAGGGG CGCCCGGTTT TTTTGTCAA GACCGACCTG 3960
 25 TCCGGTGCCC TGAATGAACT GCAGGACGAG GCAGCGCGGC TATCGTGGCT GGCCACGACG 4020
 GCGGTTCTTT GCGCAGCTGT GCTCGACGTT GTCAGTGAAG CGGGAAGGGA CTGGCTGCTA 4080
 TTGGGCGAAG TGCCGGGGCA GGATCTCCTG TCATCTCACC TTGCTCCTGC CGAGAAAGTA 4140
 30 TCCATCATGG CTGATGCAAT GCGGCGGCTG CATACTGCTG ATCCGGCTAC CTGCCCATTG 4200
 GACCACCAAG CGAAACATCG CATCGAGCGA GCACGTACTC GGATGGAAGC CGGTCTTGTC 4260
 35 GATCAGGATG ATCTGGACGA AGAGCATCAG GGGCTCGCGC CAGCCGAACT GTTCGCCAGG 4320
 CTCAAGGCGC GCATGCCCCG CGGCGAGGAT CTCGTCGTGA CCCATGGCGA TGCCTGCTTG 4380
 CCGAATATCA TGGTGGAAAA TGGCCGCTTT TCTGGATTCA TCGACTGTGG CCGGCTGGGT 4440
 40 GTGGCGGACC GCTATCAGGA CATAGCGTTG GCTACCCGTG ATATTGCTGA AGAGCTTGCC 4500
 GGCGAATGGG CTGACCGCTT CCTCGTGCTT TACGGTATCG CCGCTCCCGA TTCGCAGCGC 4560
 ATCGCCTTCT ATCGCCTTCT TGACGAGTTC TTCTGAGCGG GACTCTGGGG TTCGAAATGA 4620
 45 CCGACCAAGC GACGCCCAAC CTGCCATCAC GAGATTTGGA TTCCACCGCC GCCTTCTATG 4680
 AAAGGTTGGG CTTCCGAATC GTTTTCCGGG ACGCCGGCTG GATGATCCTC CAGCGCGGGG 4740
 50 ATCTCATGCT GGAGTTCTTC GCCCACCCCA ACTTGTTTAT TGCAGCTTAT AATGGTTACA 4800
 AATAAGCAA TAGCATCACA AATTTACAAA ATAAAGCATT TTTTCACTG CATTCTAGTT 4860

GTGGTTTGTC CAAACTCATC AATGTATCTT ATCATGTCTG TATACCGTCG ACCTCTAGCT 4920
5 AGAGCTTGCG GTAATCATGG TCATAGCTGT TTCCTGTGTG AAATTGTTAT CCGCTCACAA 4980
TTCCACACAA CACACGAGCC GGAAGCATAA AGTGTAAGC CTGGGGTGCC TAATGAGTGA 5040
GCTAACTCAC ATTAATTGCG TTGCGCTCAC TGCCCGCTTT CCAGTCGGGA AACCTGTCGT 5100
10 GCCAGCTGCA TTAATGAATC GGCCAACGCG CGGGGAGAGG CGGTTTGCCT ATTGGGCGCT 5160
CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTGCT TCGGCTGCGG CGAGCGGTAT 5220
CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA 5280
15 ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT 5340
TTTTCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA AGTCAGAGGT 5400
GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC 5460
20 GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA 5520
GCGTGGCGCT TTCTCAATGC TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTGCT 5580
CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA 5640
25 ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG 5700
GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC 5760
30 CTAACACGG CTACACTAGA AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA 5820
CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA AACCACCGCT GGTAGCGGTG 5880
GTITTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT 5940
35 TGATCTTTTC TACGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTAA GGGATTTTGG 6000
TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTAA 6060
40 AATCAATCTA AAGTATATAT GAGTAACTT GGTCTGACAG TTACCAATGC TTAATCAGTG 6120
AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTCG 6180
TGTAATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC 6240
45 GAGACCCACG CTCACCGGCT CCAGATTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG 6300
AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAAT TGTTGCCGGG 6360
AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG 6420
50 GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT 6480
CAAGGCGAGT TACATGATCC CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCTC 6540

CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTATCACT CATGGTTATG GCAGCACTGC 6600
 ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA 6660
 CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC 6720
 GGGATAATAC CGCGCCACAT AGCAGAACTT TAAAAGTGCT CATCATTGGA AAACGTTCTT 6780
 CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTGAGATC CAGTTCGATG TAACCCACTC 6840
 GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA 6900
 CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA 6960
 TACTCTTCCT TTTTCAATAT TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT 7020
 ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT TCCGCGCACA TTTCCCCGAA 7080
 AAGTGCCACC TGACGTCGAC GGATCGGG 7108

(2) INFORMATION ZU SEQ ID NO: 39:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 4019 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..4019

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 39:

CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA 60
 ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACTATGAGA 120
 GGATCGCATC ACCATCACCA TCACGGATCC CTGGTTCTGT TTGAAGGGGA CGAGGTGTAT 180
 TCCACCATCC GGAAGCAGGA ATACAATGGG AAGATCCCTC GGTTCCGCCG CATCCGGGGC 240
 GAGAGTGAGC TGTACACCAG TGATACTGTC ATGCAGAACC CACAGTTCAT CAAAGCCACC 300
 ATCGTGCACC AAGACCAGGC TTACGATGAC AAGATCTACT ACTTCTTCCG AGAGGACAAT 360
 CCTGACAAGA ATCCTGAGGC TCCTCTCAAT GTGTCCCGTG TGGCCAGTT GTGCAGGGGG 420
 GACCAGGGTG GGGAAAGTTC ACTGTCAGTC TCCAAGTGGA ACACTTTTCT GAAAGCCATG 480

| | | |
|----|--|------|
| | CTGGTATGCA GTGATGCTGC CACCAACAAG AACTTCAACA GGCTGCAAGA CGTCTTCCTG | 540 |
| 5 | CTCCCTGACC CCAGCGGCCA GTGGAGGGAC ACCAGGGTCT ATGGTGTTTT CTCCAACCCC | 600 |
| | TGGAATACT CAGCCGTCTG TGTGTATTCC CTCGGTGACA TTGACAAGGT CTTCCGTACC | 660 |
| | TCCTCACTCA AGGGCTACCA CTCAAGCCTT CCCAACCCGC GGCCTGGCAA GTGCCTCCCA | 720 |
| 10 | GACCAGCAGC CGATACCCAC AGAAAGCTTA ATTAGCTGAG CTTGGACTCC TGTTGATAGA | 780 |
| | TCCAGTAATG ACCTCAGAAC TCCATCTGGA TTTGTTTCTG ACGCTCGGTT GCCGCCGGGC | 840 |
| | GTTTTTTATT GGTGAGAATC CAAGCTAGCT TGGCGAGATT TTCAGGAGCT AAGGAAGCTA | 900 |
| 15 | AAATGGAGAA AAAAATCACT GGATATACCA CCGTTGATAT ATCCCAATGG CATCGTAAAG | 960 |
| | AACATTTTGA GGCATTTTCTG TCAGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG | 1020 |
| 20 | ATATTACGGC CTTTTTAAAG ACCGTAAAGA AAAATAAGCA CAAGTTTTAT CCGGCCTTTA | 1080 |
| | TTACATTTCT TGCCCGCCTG ATGAATGCTC ATCCGGAATT TCGTATGGCA ATGAAAGACG | 1140 |
| | GTGAGCTGGT GATATGGGAT AGTGTTCACC CTTGTTACAC CGTTTTCCAT GAGCAAACCTG | 1200 |
| 25 | AAACGTTTTT ATCGCTCTGG AGTGAATACC ACGACGATTT CCGGCAGTTT CTACACATAT | 1260 |
| | ATTCGCAAGA TGTGGCGTGT TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATG | 1320 |
| | AGAATATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT GATTTAAACG | 1380 |
| 30 | TGGCCAATAT GGACAACTTC TTCGCCCCCG TTTTACCAT GGGCAAATAT TATACGCAAG | 1440 |
| | GCGACAAGGT GCTGATGCCG CTGGCGATTG AGGTTCATCA TGCCGTCTGT GATGGCTTCC | 1500 |
| 35 | ATGTCGGCAG AATGCTTAAT GAATTACAAC AGTACTGCGA TGAGTGGCAG GGCGGGGCGT | 1560 |
| | AATTTTTTTA AGGCAGTTAT TGGTGCCCTT AAACGCCTGG GGTAATGACT CTCTAGCTTG | 1620 |
| | AGGCATCAAA TAAACGAAA GGCTCAGTCG AAAGACTGGG CCTTTCGTTT TATCTGTTGT | 1680 |
| 40 | TTGTCGGTGA ACGCTCTCCT GAGTAGGACA AATCCGCCGC TCTAGAGCTG CCTCGCGCGT | 1740 |
| | TTCCGTGATG ACGGTGAAAA CCTCTGACAC ATGCAGCTCC CGGAGACGGT CACAGCTTGT | 1800 |
| 45 | CTGTAAGCGG ATGCCGGGAG CAGACAAGCC CGTCAGGGCG CGTCAGCGGG TGTGGCGGG | 1860 |
| | TGTCGGGGCG CAGCCATGAC CCAGTCACGT AGCGATAGCG GAGTGTATAC TGGCTTAACT | 1920 |
| | ATGCGGCATC AGAGCAGATT GTACTGAGAG TGCACCATAT GCGGTGTGAA ATACCGCACA | 1980 |
| 50 | GATGCGTAAG GAGAAAATAC CGCATCAGGC GCTCTTCCGC TTCCTCGCTC ACTGACTCGC | 2040 |
| | TGCGCTCGGT CTGTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGCG GTAATACGGT | 2100 |

TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG 2160
 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCTGACG 2220
 5 AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT 2280
 ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCCTCTCC GTTCCGACC CTGCCGCTTA 2340
 10 CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCAGCT 2400
 GTAGGTATCT CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC 2460
 CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATCG TCTTGAGTCC AACCCGTAA 2520
 15 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG 2580
 TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG 2640
 TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT 2700
 20 GATCCGGCAA ACAAAACCACC GCTGGTAGCG GTGGTTTTTT TGTGCAAG CAGCAGATTA 2760
 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGG TCTGACGCTC 2820
 AGTGGAACGA AAACCTACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA 2880
 25 CCTAGATCCT TTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA 2940
 CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT 3000
 30 TTCGTTTCATC CATAGCTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT 3060
 TACCATCTGG CCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTACCG GCTCCAGATT 3120
 TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT 3180
 35 CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGAAGCTAG AGTAAGTAGT TCGCCAGTTA 3240
 ATAGTTTGC CAACGTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTG 3300
 GTATGGCTTC ATTCAGCTCC GGTTCCTAAC GATCAAGGCG AGTTACATGA TCCCCATGT 3360
 40 TGTGCAAAAA AGCGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG 3420
 CAGTGTATC ACTCATGGT ATGGCAGCAC TGCATAATC TCTTACTGTC ATGCCATCCG 3480
 TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC 3540
 45 GCGGACCGAG TTGCTCTTG CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA 3600
 CTTTAAAAGT GCTCATCATT GGAACCGTT CTCGGGGCG AAAACTCTCA AGGATCTTAC 3660
 50 CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT 3720
 TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAAGG 3780

GAATAAGGGC GACACGGAAA TGTGAATAC TCATACTCTT CCTTTTCAA TATTATTGAA 3840
 GCATTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA 3900
 5 AACAAATAGG GGTTCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC TAAGAAACCA 3960
 TTATTATCAT GACATTAACC TATAAAAATA GCGTATCAC GAGGCCCTTT CGTCTTCAC 4019

10

(2) INFORMATION ZU SEQ ID NO: 40:

(i) SEQUENZ CHARAKTERISTIKA:

15

- (A) LÄNGE: 3999 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: DNS (genomisch)

20

(ix) MERKMALE:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..3999

25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 40:

CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA 60
 30 ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACTATGAGA 120
 GGATCGCATC ACCATCACCA TCACACGGAT CCGCATGCGA GCTCCCAGTG GGAGGTGAGC 180
 CAGGTGCCCC TGGACCTGTG TGAGGTCTAT GCGGGGGCT GCCACGGTTG CCTCATGTCC 240
 35 CGAGACCCCT ACTGCGGCTG GGACCAGGGC CGCTGCATCT CCATCTACAG CTCCGAACGG 300
 TCAGTGCTGC AATCCATTAA TCCAGCCGAG CCACACAAGG AGTGTCCCAA CCCCACCA 360
 GACAAGGCCC CACTGCAGAA GGTTTCCCTG GCCCCTACTA CTCGCTACTA CCTGAGCTGC 420
 40 CCCATGGAAT CCCGCCACGC CACCTACTCA TGGCGCCACA AGGAGAACGT GGAGCAGAGC 480
 TGCGAACCTG GTCACCAGAG CCCCCTACTG ATCTGTTC TCGAGAACCT CACGGCGCAG 540
 45 CAGTACGGCC ACTACTTCTG CGAGGCCAG GAGGGCTCCT ACTCCGCGA GGCTCAGCAC 600
 TGGCAGCTGC TGCCCGAGGA CGGCATCATG GCCGAGCACC TGCTGGGTCA TGCCTGTGCC 660
 CTGGCTGCCT CCCTCTGGCT GGGGTGCTG CCCCACTCA CTCTTGCTT GCTGGTCCAC 720
 50 GTGAAGCTTA ATTAGCTGAG CTTGGACTCC TGTGATAGA TCCAGTAATG ACCTCAGAAC 780
 TCCATCTGGA TTTGTTTACA ACGCTCGGTT GCCGCCGGC GTTTTTTATT GGTGAGAATC 840

55

CAAGCTAGCT TGGCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA AAAAATCACT 900
 GGATATACCA CCGTTGATAT ATCCCAATGG CATCGTAAAG AACATTTTGA GGCATTTTCAG 960
 5 TCAGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG ATATTACGGC CTTTTTAAAG 1020
 ACCGTAAAGA AAAATAAGCA CAAGTTTTAT CCGGCCTTTA TTCACATTCT TGCCCGCCTG 1080
 10 ATGAATGCTC ATCCGGAATT TCGTATGGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT 1140
 AGTGTTCAAC CTTGTTACAC CGTTTTCCAT GAGCAAACCTG AAACGTTTTTTC ATCGCTCTGG 1200
 AGTGAATACC ACGACGATTT CCGGCAGTTT CTACACATAT ATTCGCAAGA TGTGGCGTGT 1260
 15 TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG AGAATATGTT TTTCGTCTCA 1320
 GCCAATCCCT GGGTGAGTTT CACCAGTTTT GATTTAAACG TGGCCAATAT GGACAACTTC 1380
 20 TTCGCCCCCG TTTTCACCAT GGGCAAATAT TATACGCAAG GCGACAAGGT GCTGATGCCG 1440
 CTGGCGATTC AGGTTTCATCA TGCCGTCTGT GATGGCTTCC ATGTCGGCAG AATGCTTAAT 1500
 GAATTACAAC AGTACTGCGA TGAGTGGCAG GGCGGGCGT AATTTTTTTA AGGCAGTTAT 1560
 25 TGGTGCCCTT AAACGCCTGG GGTAACTGACT CTCTAGCTTG AGGCATCAAA TAAAACGAAA 1620
 GGCTCAGTCG AAAGACTGGG CCTTTCGTTT TATCTGTTGT TTGTCGGTGA ACGCTCTCCT 1680
 GAGTAGGACA AATCCGCCGC TCTAGAGCTG CCTCGCGCGT TTCGGTGATG ACGGTGAAAA 1740
 30 CCTCTGACAC ATGCAGCTCC CGGAGACGGT CACAGCTTGT CTGTAAGCGG ATGCCGGGAG 1800
 CAGACAAGCC CGTCAGGGCG CGTCAGCGGG TGTTGGCGGG TGTCGGGGCG CAGCCATGAC 1860
 35 CCAGTCACGT AGCGATAGCG GAGTGTATAC TGGCTTAACT ATGCGGCATC AGAGCAGATT 1920
 GTACTGAGAG TGACCATAT GCGGTGTGAA ATACCGCACA GATGCGTAAG GAGAAAATAC 1980
 CGCATCAGGC GCTCTCCGC TTCCTCGCTC ACTGACTCGC TGCGCTCGGT CTGTCGGCTG 2040
 40 CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT 2100
 AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC 2160
 GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCTGACG AGCATCACAA AAATCGACGC 2220
 45 TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA 2280
 AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT 2340
 50 CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCACGCT GTAGGTATCT CAGTTCGGTG 2400
 TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC 2460

| | | |
|----|--|------|
| | GCCTTATCCG GTAACATCG TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG | 2520 |
| 5 | GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC | 2580 |
| | TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG | 2640 |
| | CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC | 2700 |
| 10 | GCTGGTAGCG GTGGTTTTTT TGTTCGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT | 2760 |
| | CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGAACGA AAACCTCACGT | 2820 |
| | TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT TTAAATTAA | 2880 |
| 15 | AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA | 2940 |
| | TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTTCATC CATAGCTGCC | 3000 |
| 20 | TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT | 3060 |
| | GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA | 3120 |
| | GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT | 3180 |
| 25 | AATTGTTGCC GGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT | 3240 |
| | GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC | 3300 |
| 30 | GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCATGT TGTGCAAAAA AGCGGTTAGC | 3360 |
| | TCCTTCGGTC CTCGATCGT TGTCAGAAGT AAGTTGGCCG CAGTGTATC ACTCATGGTT | 3420 |
| | ATGGCAGCAC TGCATAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT | 3480 |
| 35 | GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC GCGACCGAG TTGCTCTTGC | 3540 |
| | CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT | 3600 |
| 40 | GGAAAACGTT CTTGCGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG | 3660 |
| | ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTTCT | 3720 |
| | GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA | 3780 |
| 45 | TGTTGAATAC TCATACTCTT CCTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT | 3840 |
| | CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCGCGC | 3900 |
| 50 | ACATTTCCCC GAAAAGTGCC ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC | 3960 |
| | TATAAAAATA GGCGTATCAC GAGGCCCTTT CGTCTTCAC | 3999 |

(2) ANGABEN ZU SEQ ID NO: 41:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 8888 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: exon
- (B) LÄNGE: 1..8888

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 41:

| | | | | | | |
|------------|-------------|------------|------------|------------|------------|------|
| GAGCCGCACA | CGGTGCTTTT | CCACGAGCCA | GGCAGCTCCT | CTGTGTGGGT | GGGAGGACGT | 60 |
| GGCAAGGTCT | ACCTCTTTGA | CTTCCCCGAG | GGCAAGAACG | CATCTGTGCG | CACGGTGAGC | 120 |
| CTCTCTCTTC | CCCCAACACC | CCCCCTACCC | TCTTATCTCC | CCTCTGGCCC | TGCCAAGGGT | 180 |
| CCTCAGGGAA | TCCGAGGGAG | CTGGCTTCTC | TTCTTAAACT | GCCCCACCT | CCGTATCCTA | 240 |
| TAAATGGCTC | CTGGGGGAGG | CTCCCTAAAG | GTAGTCCAGA | TTGGAGTGGG | GAGCTGGGGC | 300 |
| GGTGTGGAGA | AAAACAGGAG | CTAATGGGCC | TGGCCAGCTG | GGCAGCGCTG | CTGCGGAAAG | 360 |
| CCCAGGCTGG | AAGCTGGGCC | CCAGAGCCCA | TGCCTGGTCT | TCTGAACCCT | CTGGGCCTCA | 420 |
| GCTCTGGATA | TGAGACCCTG | TTTGACCTCA | GGTAGATCAC | TCACCCTCTC | AGAGCCCCAG | 480 |
| TTGCTCATCT | GTCAGATGAG | AATAATGGTT | GCTTCCTTTG | GGGCTTATCC | TGAGGCTGTG | 540 |
| TGGAAAGCAT | TTCAGGGGTA | CCTCACCCCT | GGCAGATTGA | ACTAATGCTT | CTCCCCITCC | 600 |
| CCAGGTGAAT | ATCGGCTCCA | CAAAGGGGTC | CTGTCTGGAT | AAGCGGGTGA | GCGGGGGAGG | 660 |
| GATCTGGAGG | GGTCTGAGCC | ACTTGGTAAA | GGGAGAGGAG | ACCCTGAGGG | TCTAAGGAAG | 720 |
| GAAGCATGGC | CCTGCCCCAC | GAGTCCCAGA | CTGATGGGGA | GACGTGGTCC | TCTGTGCTTA | 780 |
| GGGGATGGCG | TCAGCTGCAC | ACACTCTGGG | CTGTCCCGGG | AGGCTGTCAC | CTATGCTAAG | 840 |
| CCCTTCTGAC | ACCTTCTTCC | CTGATCCTGG | GGGTCCTAGT | GCTAGGCTTG | CCAGGGCCTT | 900 |
| CCAGCAACCA | ATTTCTCTCC | TCCCTTCTCT | CTTCCCCGGG | CAGGACTGCG | AGAACTACAT | 960 |
| CACTCTCCTG | GAGAGGCGGA | GTGAGGGGCT | GCTGGCCTGT | GGCACCAACG | CCCGGCACCC | 1020 |
| CAGCTGCTGG | AACCTGGTGA | GAAGGCTGCT | CCCCATGTGC | CTGATCAGCT | CACCTTCTAC | 1080 |
| TGCGTGGGCT | TCTGCCCCCTC | ATGGTGGGAA | GGAGATGGCG | AGACTCCAAT | GCTGGCCTTG | 1140 |

CCCTGGGAGG ATGGGGCTCC TGGCCGAGAA ACTGGCCGTC ATGGGAGGCA GTGGCTGTGG 1200
 GATTATGTGG CCATCCAACC CTCTGGATCT CCCACAGGTG AATGGCACTG TGGTGCCACT 1260
 5 TGGCGAGATG AGAGGCTACG CCCCCCTCAG CCCGGACGAG AACTCCCTGG TTCTGTTTGA 1320
 AGGTTGGGGC ATGCTTCGGA ACTGGGCTGG GAGCAGGATG GTCAGCTCTT TGTCCAGTGT 1380
 10 CCGGAGGAGG GACTTCCAGG AGCTGCCTGC CCTTACTCAT TTCTCCCTCC CACTGACCCC 1440
 AGGGGACGAG GTGTATTCCA CCATCCGGAA GCAGGAATAC AATGGGAAGA TCCCTCGGTT 1500
 CCGCCGCATC CGGGGCGAGA GTGAGCTGTA CACCAGTGAT ACTGTCATGC AGAGTGAGTC 1560
 15 AGGCTCCGGC TGGGCTGAGG GTGGGCAAGG GGGTGTGAGC ACTTAAGGTG GCAGATGGGA 1620
 TCCTGATGTT TCTGGGAGGG CTCCCTGAGG GCCGCTGGGG CCATGCAGGA AAGCAGGACC 1680
 TTGGTATAGG CCTGAGAAGT TAGGGTTGGC TGGGAGCAGA GGAACAGACA AGGTATAGCA 1740
 20 GTGGGATGGG CCCAGCCCTC TTCAGGAACA CAAACAGAGG GAGCCCCAGA CCCAGTGACG 1800
 GGTCCCCAGG AGCCAAAGTT TATCCTCTGC TGAGTTCACG TGGAGGCAGC CCCCCAACTC 1860
 CCTCCTCATC AGGGCTCTGC CAATTGAGCA GAAGTGACAT AGGGGCCCCC AGGGACCTTC 1920
 25 CCCCCTCCC CAGGCATGAA GTCATTGCTC CTGGGCCGAT GACATCTTTG TAGGAAGAGG 1980
 GCAAAACAGG TGTGGGGTGG AGGTGCAGGG TCTAGGGCCC CTCGGGGAGT TGGACCTGAT 2040
 30 GTTATGAGTC CTATTCCAGA TCTGATTGTC CATGGTTTGT GCAGACCGA AGGAGGGAGG 2100
 AGAGTGTGCA GGGTTGGAAT GGTCTCCCGG GCAAGCTTCC CAGCCTTACG CCCATTCGCT 2160
 TCTGTGCCCT GGCAGACCCA CAGTTCATCA AAGCCACCAT CGTGCACCAA GACCAGGCTT 2220
 35 ACGATGACAA GATCTACTAC TTCTCCGAG AGGACAATCC TGACAAGAAT CCTGAGGCTC 2280
 CTCTCAATGT GTCCCGTGTG GCCCAGTTGT GCAGGGTGAA CACGGGCGTG AGGGCTGCTG 2340
 GCTACGTGTC TGTGCATGAA TAGGCCTGAG TGAGGGTGAG TTCTGTGTGT CCGTGTGCAT 2400
 40 GTAGAAGTTG TGTGGATGTA TGAGTGGGTC TGTGTCAGGG ACTGTGGGAG CAGCTGTGTG 2460
 TGCATGGAGC ATCATGTGTC TGTGTGTGGG TAAAGGTGGC TGAGCTCCTG TGCACGTATG 2520
 45 ATGGCGTGTG AGCGTGTGTA TGATGGGGTG TGTGTGTGTG TGTGTGTGTG TGTTTGCCT 2580
 GTGTGAATGT GCTGTGCCAC GTATGTGGGT GCGTGAGTCA GTAAATGTGT GTCTGAGTCC 2640
 GTCTGCTCTG TGGGGACCTG GCACTCTCAC CTGCCCTGAC CCTGGGCACT GCTGGCCCTG 2700
 50 GGCTCTGGAT CAGCCAGGCC TGCTTGACAG AGTCTCATCT GGAGACCTGC CCTGAGTCCT 2760
 GGGGCACCCC CGGCAGGTCC TGGCCCCTCG CAGCCTGCCT TCCTCCTCTG GGCCAGGTG 2820

TTGATATTGC TGGCAGTGGT TTCCTGGGGT GTGTGGGGAA GCCCCGGCAG GTGCTGAGGG 2880
 GCCTCTTCTC CCCTCTACCC TTCCAGGGGG ACCAGGGTGG GGAAAGTTCA CTGTCACTCT 2940
 5 CCAAGTGGAA CACTTTTCTG AAAGCCATGC TGGTATGCAG TGATGCTGCC ACCAACAAGA 3000
 ACTTCAACAG GCTGCAAGAC GTCTTCCTGC TCCCTGACCC CAGCGGCCAG TGGAGGGACA 3060
 10 CCAGGGTCTA TGGTGTTTTT TCCAACCCCT GGTGAGTGGC CCTTGTCTTG GGGCCGGGGC 3120
 TGGCATTGGT TCAGTGTTCA GTAGGGACAG GAGGCCTTGG GCCCTGCTGA GGGCCTCCCT 3180
 GGTGTGGCAG GAGCAGGGGC TGCAGGCTCA AGAGGCTGGG CTGTGCTGG GTGTGGGGTG 3240
 15 GGGGGACAGC CAGTGCGATG TATGTACTGT TGTGTGAGTG AGTCTGCACT CATGGGTGTG 3300
 TGTGCATGCC CTATATGCAC ACTCATGACT GCACTTGTGC CTGTGTGTCC CACCACCTGC 3360
 TTGTGCCGAG AGTGGACACT GGGCCCAGGA GGAAGCTGCT GAAGCATCTC TCGGGGAGCT 3420
 20 GGGTGCTATT ACACCTGCTC AGGCACTGCC TGAGCCCGAT AATTCACACT TCTTAATCAC 3480
 TCTCATTGAT TGAACACACG GCAGGCGGAA GTGTTGGGTG TGTGTGGGA GAGTTAGGA 3540
 25 TAGAGTGGAG GAAGCCAAGA CCCTGCTCTG TGGCTCCTGG GTGAGTGGGT CCCCCAGGCT 3600
 GGGAAAGGGT TGGGGGTCTG GCCTCCTGGG GCATCAGCAC CCCACAGCCT GTGCCCAGGG 3660
 AGGGCTAGAG AACTGCTCAG CCTATGATGG GGTTCCTCCT GCCTTGGGGT TGGGTAGAGC 3720
 30 AGATGGCCTC TAGACTCAGT GATTCTGTAA CAGGATACAA GTTTGTGGTT TTAAATTGCA 3780
 GCACAAAGAA ATTAGGCTGA ACTCCTCTCC TTCCTCCTCT CCATCCCTCC CCATTTTCAG 3840
 TGGTGGTTGG CAACTCAGTG CCAGGCACAA GGCTGGCCTG GGTGAGTGGA GGTGGATGGG 3900
 35 TGGGTTCTGG GCCCCCATT GAGCTGGTCT CCATGTCACT GCAGGAACTA CTCAGCCGTC 3960
 TGTGTGTATT CCCTCGGTGA CATTGACAAG GTCTTCCGTA CCTCCTCACT CAAGGGCTAC 4020
 40 CACTCAAGCC TTCCCAACCC GCGGCCTGGC AAGGTGAGCG TGACACCAGC CGTGGCCCAG 4080
 GCCCAGCCCT CTTCTGCCT CACCTCCCAC CACCCCACTG ACCTGGGCCT GCTCTCCTTG 4140
 CCCAGTGCCT CCCAGACCAG CAGCCGATAC CCACAGAGAC CTTCCAGGTG GCTGACCGTC 4200
 45 ACCCAGAGGT GGCGCAGAGG GTGGAGCCCA TGGGGCCTCT GAAGACGCCA TTGTTCCACT 4260
 CTAAATACCA CTACCAGAAA GTGGCCGTCC ACCGCATGCA AGCCAGCCAC GGGGAGACCT 4320
 TTCATGTGCT TTACCTAACT ACAGGTGAGA GGCTACCCCG GGACCCTCAG TTTGCTTTGT 4380
 50 AAAACGGGC ATGAAAGGTG TAAGGAATAA TGTAGTTAAC ATCTGGTTGG ATCTTTACAT 4440

| | | |
|----|--|------|
| | GTGGAAGGAA TAATTGAGTG ACTGGAGTTG TCAGGGGTTA ATGTGTGTGG GTGTGGAAGA | 4500 |
| 5 | GCCAGGCAGG GAGAGCTTCC TGGAGGAGGT AGGGGCAAGA GGGAAAGGGG GATGGGAGAA | 4560 |
| | AAGCAAGCAC TGGGATTTGG AGGCGGAAAT CTGGAGAGTC TGAGCAAAGC CAGGTGCACC | 4620 |
| | TTTGGTCCAG ATGTCTGACT CAGGGAAGAA GATGGTAGGA AGAGACGTGG CAAATGAGGA | 4680 |
| 10 | GGAGGGGCCT GAACCACAGG GATACTGGCC TCTGCCAGGC AGAATGAGGG AGTCAGGCCC | 4740 |
| | TGCGCCTGTC TTTGGGATTG TGCAGGTGAG AAGAAACATT TGAGGAGTTG ATGGGGCACA | 4800 |
| | AATTAGGTAT GGGGAAGGAG TTCCAGGGGG CAGAACCTTT GCCATCTCAC AGAGGACAGG | 4860 |
| 15 | GGCAGCTTCT CTTCTTCCCT GGAGTAGGCC CTGCTGGGGG AAGCTGGGTG GAATGCCGTG | 4920 |
| | GGAGATGCTC CTGCTTCTG GAAAGCCACA GGACACGGAG GAGCCAGTCC TGAGTTGGGT | 4980 |
| 20 | TTGTGCGCAGC TTCCCATGCC AGCTGCCTTC CTTGAGACTG GAAAGGGCCT CTAGCACCCC | 5040 |
| | TGGGGCCATT CAATTCAGGC CCAGGCGCCC AACCTCAGTT GTTCACATTC CCCATGTGAT | 5100 |
| | CTCCTGTTGC TGCTTCACCT TGGGACTGTC TCGGCTTTGG TGACCTTGTA GGAAACTGGA | 5160 |
| 25 | ACCCCAGCAC CATTGTTTGG CTCCTGGAAG CCTTGGGGAG AGGAATTTCC CACAGGGCAG | 5220 |
| | GGCCTGGGTC CTGATTCCCT GCCTCTTTAC TCCCTATTCA TCCCGGCTAC ACCCTTGGGC | 5280 |
| | CCCCATCCTT GCTTGGCTCC AGTACTGGCT GGCACAGCTG TTGTGGTCAT CCAGGGATGG | 5340 |
| 30 | CAGGGCACTG GGGAACAGAA GAGAGAGGTC ACACAGTGCG GAACTGGGAG CAGGAGCTAG | 5400 |
| | GACAAGGAAG GCTGGACTTG GGCCATGGAT TCCCTTCCTG CAGACTTGGG AAGTGAGCAC | 5460 |
| 35 | ACTTGAGTGA TTAGAGAAGG TGTCTTCGTT CTAAGGGCAG TGGAGGAGGC ACCATTTTGG | 5520 |
| | AGCCTGCATC ATTCTGATTT GGGCTAGATT GAAAAATAGA GCTTTCTAAG TCCTCTGCAG | 5580 |
| | AGAATGGGAG GCTCTCACAA CTGGGAGAAG TATTGGCTCT TTTCCTGAGA ATTTTGCCAA | 5640 |
| 40 | GGGTATGCTG TTA CTGGGGC TGGTTTGGA GAGTATAGG GCATTATGTC TGTGAAGGCA | 5700 |
| | GTGGCTGGGG TGGGGCCTTA TCAGGCCCAA GGAGCATCTG GCCACATCTC AGAGTCCACA | 5760 |
| | GATGAGGATC ACGGATGTGT AGAGGAAACA TCCTAGGCAG GCAATCATCT GACTGCTTTT | 5820 |
| 45 | TTGGGGCAGG TGATGCCCTG GGAAATTGGG AGGGAGGGAG AGAGGGAGGT AGGCTATTCT | 5880 |
| | AGAAACTGGG AGAGCAGGTG AGGTAGGATT GGGAGGACCA GGGGTCAGGG TCCCCATTGG | 5940 |
| 50 | TCCCTAATTG AGAACGGAGA GAGCATTGGT CTAGGAGGCA GGCAGCTCGG TTATAAGACC | 6000 |
| | TTGGGAACTC TTGATTTAGA ATCCAAGATC CTTTTTAGAT CTAGGATTTT ATAAAATTAA | 6060 |
| 55 | GATATCCCCT AAGATCAAAT GCAACGTGGA GTCCTGAATT GGATCCTAGA ACAGAAGAAG | 6120 |

.TTTGTG GAAAACTAG TGAAATCCAA ATAAAGTCTG TAGTTTTGTT AATAGTAATG 6180
 ACCAATGTC AGTTGCCTAG TTGTGACAAA TATACCGTGG TTATGTAAGA TGGTAACATT 6240
 AGGGGGAAGT GGAGAAGGGT AGATTGGAGC TCTCTGTACT ATCTTTGCAA CTTTTCTGGG 6300
 AATCTAAAAT TACTCCAAAA TAAAAAATAA ATGTATTATA AGTAAATATA TTCCTAAGA 6360
 GTCCAGGAGG CAGGGGAGTT GTAGAAGCAG CTGAGTGGTT GGGTTCTGAC AGATTTGGTT 6420
 CCAACTCGGT CTCTGCTGCT CACCAGCTGT GTGACCTTGA GCAAGTGGCT TAGCCTTTCT 6480
 GAGCCTGATT TCCTTATCTG TGGAGTGGGG AAGATGACAG CCACCTCGCA GGGCTGTGGA 6540
 GGGTTAAACG AGGTGATGCA TGGACAGCAG CCGCACTGAC CTTGCTGGTG TGGGGCTCCT 6600
 GCTTCTGTTT TTCCCGTGCA GCCTTGGGAA TGTGAGGAGC CGTATCCAGG GACCCCTGGG 6660
 CCTCCTGGGA TGGCCTCTCT GGATCAGCCT TGGAAAGTTC CAGGCTGCCC TTAGGCTCCC 6720
 ACATTCTTCC CCAGTCACGC TCTCCTCGCC CTGCCCCACAC CAGTCCTGTG ACCCTTGCCT 6780
 GAGTTGTGAC TTCCACCCCC TCCCCGGCCT AGAGGAAAGC TGCCTGGCCC CTCAGTGGGA 6840
 CTCCCGCCCA CTGACCCTCT GTCCACCATA CACAGACAGG GGCATATCC ACAAGGTGGT 6900
 GGAACCGGGG GAGCAGGAGC ACAGCTTCGC CTTCAACATC ATGGAGATCC AGCCCTTCCG 6960
 CCGCGCGGCT GCCATCCAGA CCATGTCGCT GGATGCTGAG CGGGTGAGCC TTCCCCACT 7020
 GCGTCCCATG GGCTATGCAG TGAATGCAGC TGAGGACAGG GCTCCTTTGC ATGTGATTTG 7080
 TGTGTTCTTT TAAGAGCTTC TAGGCCTTAG GGCCTGGACA TTTAGGACTG AGTGTGGGGT 7140
 GGGGCCCCGG CCTGACCCAA TCCTGCTGTC CTTCCAGAGG AAGCTGTATG TGAGCTCCCA 7200
 GTGGGAGGTG AGCCAGGTGC CCCTGGACCT GTGTGAGGTC TATGGCGGGG GCTGCCACGG 7260
 TTGCCTCATG TCCCGAGACC CTTACTGCGG CTGGGACCAG GGCCGCTGCA TCTCCATCTA 7320
 CAGCTCCGAA CGGTACGTTG GCCGGGATCC CTCCGTCCCT GGGACAAGGT GGGCATGGGA 7380
 CAGGGGGAGG TGTGTCGGG CTGGAAGAGG TGGCGTACT GGGCCTTTCT TGTGGGACCT 7440
 CCTCTCTACT GGAATGAC TAGGGGTAAG GATATGAGG TCAGGTCTGC AGCCTTGTAT 7500
 CTGCTGATCC TCTTTCGTCC TTCCCACTCC AGGTCAGTGC TGCAATCCAT TAATCCAGCC 7560
 GAGCCACACA AGGAGTGTCC CAACCCAAA CCAGGTACCT GATCTGGCCC TGCTGGCGGC 7620
 TGTGGCCCAA TGAGTGGGGT ACTGCCCTGC CTGATTGTC CTGGTCTGAG GGAAACATGG 7680
 CCTGTCTCTG TGGGCCCCAG GTACATGGGG CAGGATACAG TCCTGCAGAG GGAGCCCTCT 7740

TGGTGGGATG AGCGAGACGG GAGAAAAAAG GAGGACGCTG AGGGCTGGGT TCCCCACGTT 7800
 CATTGAGAAG CCTTGTCTTG GGATCCCAGT CGGTGGGGAG GACACATCCT CCCCTGGGAG 7860
 CTCTTTGTCC CTCCTCACGG CTGCTTCCCC ACTGCCTCCC CAGACAAGGC CCCACTGCAG 7920
 AAGGTTTCCC TGGCCCCAAA CTCTCGCTAC TACCTGAGCT GCCCCATGGA ATCCCGCCAC 7980
 GCCACCTACT CATGGCGCCA CAAGGAGAAC GTGGAGCAGA GCTGCGAACC TGGTCACCAG 8040
 AGCCCCAACT GCATCCTGTT CATCGAGAAC CTCACGGCGC AGCAGTACGG CCACTACTTC 8100
 TGCAGGCCC AGGAGGGCTC TACTTCCGC GAGGCTCAGC ACTGGCAGCT GCTGCCCCGAG 8160
 GACGGCATCA TGGCCGAGCA CCTGCTGGGT CATGCCTGTG CCCTGGCCGC CTCCCTCTGG 8220
 CTGGGGGTGC TGCCCACT CACTCTTGGC TTGTGGTCC ACTAGGCCT CCCGAGGCTG 8280
 GGCATGCCTC AGGCTTCTGC AGCCCAGGGC ACTAGAAGT CTCACACTCA GAGCCGGCTG 8340
 GCCCGGAGC TCCTTGCCTG CCACTTCTTC CAGGGGACAG AATAACCCAG TGGAGGATGC 8400
 CAGGCCTGGA GACGTCCAGC CGCAGGCGGC TGCTGGGCCC CAGGTGGCGC ACGGATGGTG 8460
 AGGGGCTGAG AATGAGGGCA CCGACTGTGA AGCTGGGGCA TCGATGACCC AAGACTTTAT 8520
 CTCTTGAAA ATATTTTCA GACTCCTCAA ACTTGACTAA ATGCAGCGAT GCTCCCAGCC 8580
 CAAGAGCCCA TGGGTCGGG AGTGGGTTG GATAGGAGAG CTGGGACTCC ATCTCGACCC 8640
 TGGGGCTGAG GCCTGAGTCC TTCTGGAATC TTGGTACCCA CATTGCCTCC TTCCCCTCCC 8700
 TCTCTCATGG CTGGGTGGCT GGTGTTCTG AAGACCCAGG GCTACCCTCT GTCCAGCCCT 8760
 GTCCTCTGCA GCTCCCTCTC TGGTCTGGG TCCCACAGGA CAGCCGCCTT GCATGTTTAT 8820
 TGAAGGATGT TTGCTTCCG GACGGAAGGA CGAAAAAGC TCTGAAAAA AAAAAAAAAA 8880
 AAAAAAA 8888

(2) ANGABEN ZU SEQ ID NO: 42:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 6622 Basenpaare

(B) ART: Nucleotid

(C) STRANGFORM: Einzelstrang

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: exon

(B) LÄNGE: 1..6622

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 42:

| | | |
|----|--|------|
| 5 | GATATCATGG AGATAATTAA AATGATAACC ATCTCGCAA TAAATAAGTA TTTTACTGTT | 60 |
| | TTCGTAACAG TTTTGTAAATA AAAAAACCTA TAAATATGAA ATTCTTAGTC AACGTTGCCC | 120 |
| | TTGTTTTTAT GGTGCTATAC ATTTCTTACA TCTATGCGGA TCGATGGGGA TCCGCCCAGG | 180 |
| 10 | GCCACCTAAG GAGCGGACCC CGCATCTTCG CCGTCTGGAA AGGCCATGTA GGGCAGGACC | 240 |
| | GGGTGGACTT TGGCCAGACT GAGCCGCACA CGGTGCTTTT CCACGAGCCA GGCAGCTCCT | 300 |
| | CTGTGTGGGT GGGAGGACGT GGCAAGGTCT ACCTCTTTGA CTTCCCCGAG GGCAAGAACG | 360 |
| 15 | CATCTGTGCG CACGGTGAAT ATCGGCTCCA CAAAGGGGTC CTGTCTGGAT AAGCGGGACT | 420 |
| | GCGAGAACTA CATCACTCTC CTGGAGAGGC GGAGTGAGGG GCTGCTGGCC TGTGGCACCA | 480 |
| | ACGCCCCGCA CCCCAGCTGC TGAACCTGG TGAATGGCAC TGTGGTGCCA CTTGGCGAGA | 540 |
| 20 | TGAGAGGCTA TGCCCCCTTC AGCCCGGACG AGAACTCCCT GGTCTGTTT GAAGGGGACG | 600 |
| | AGGTGTATTC CACCATCCGG AAGCAGGAAT ACAATGGGAA GATCCCTCGG TTCCGCCGCA | 660 |
| 25 | TCCGGGGCGA GAGTGAGCTG TACACCAGTG ATRACTGTCAT GCAGAACCCA CAGTTCATCA | 720 |
| | AAGCCACCAT CGTGACCAA GACCAGGCTT ACGATGACAA GATCTACTAC TTCTTCCGAG | 780 |
| | AGGACAATCC TGACAAGAAT CCTGAGGCTC CTCTCAATGT GTCCCGTGTG GCCCAGTTGT | 840 |
| 30 | GCAGGGGGGA CCAGGTGGG GAAAGTTCAC TGTCAGTCTC CAAGTGGAAC ACTTTTCTGA | 900 |
| | AAGCCATGCT GGTATGCAGT GATGTGCCA CCAACAAGAA CTCAACAGG CTGCAAGACG | 960 |
| 35 | TCTTCCTGCT CCCTGACCCC AGCGGCCAGT GGAGGGACAC CAGGGTCTAT GGTGTTTTCT | 1020 |
| | CCAACCCTG GAACTACTCA GCCGTCTGTG TGTATTCCCT CGGTGACATT GACAAGGTCT | 1080 |
| | TCCGTACCTC CTCACTCAAG GGCTACCACT CAAGCCTTCC CAACCCGCGG CCTGGCAAGT | 1140 |
| 40 | GCCTCCCAGA CCAGCAGCCG ATACCCACAG AGACCTTCCA GGTGGCTGAC CGTCACCCAG | 1200 |
| | AGGTGGCGCA GAGGGTGGAG CCCATGGGGC CTCTGAAGAC GCCATTGTTC CACTCTAAAT | 1260 |
| | ACCACTACCA GAAAGTGGCC GTTCACCGCA TGCAAGCCAG CCACGGGGAG ACCTTTCATG | 1320 |
| 45 | TGCTTTACCT AACTACAGAC AGGGGCACTA TCCACAAGGT GGTGGAACCG GGGGAGCAGG | 1380 |
| | AGCACAGCTT CGCCTTCAAC ATCATGGAGA TCCAGCCCTT CCGCCGCGG GCTGCCATCC | 1440 |
| 50 | AGACCATGTC GCTGGATGCT GAGCGGAGGA AGCTGTATGT GAGCTCCCAG TGGGAGGTGA | 1500 |
| | GCCAGGTGCC CCTGGACCTG TGTGAGGTCT ATGGCGGGG CTGCCACGGT TGCCTCATGT | 1560 |
| 55 | | |

| | | |
|----|--|------|
| | CCCCGAGACCC CTACTGCGGC TGGGACCAGG GCCGCTGCAT CTCCATCTAC AGCTCCGAAC | 1620 |
| | GGTCAGTGCT GCAATCCATT AATCCAGCCG AGCCACACAA GGAGTGTCCT AACCCTCAAAC | 1680 |
| 5 | CAGACAAGGC CCCACTGCAG AAGGTTTCCC TGGCCCCAAA CTCTCGCTAC TACCTGAGCT | 1740 |
| | GCCCCATGGA ATCCCGCCAC GCCACCTACT CATGGCGCCA CAAGGAGAAC GTGGAGCAGA | 1800 |
| 10 | GCTGCGAACC TGGTCACCAG AGCCCCAACT GCATCCTGTT CATCGAGAAC CTCACGGCGC | 1860 |
| | AGCAGTACGG CCACTACTTC TGCGAGGCCC AGGAGGGCTC CTACTTCCGC GAGGCTCAGC | 1920 |
| | ACTGGCAGCT GCTGCCCCGAG GACGGCATCA TGGCCGAGCA CCTGCTGGGT CATGCCTGTG | 1980 |
| 15 | CCCTGGCTGC CTGAATTCGA AGCTTGGAGT CGACTCTGCT GAAGAGGAGG AAATTCTCCT | 2040 |
| | TGAAGTTTCC CTGGTGTTC AAGTAAAGGA GTTTGCACCA GACGCACCTC TGTTCACTGG | 2100 |
| 20 | TCCGGCGTAT TAAAACACGA TACATTGTTA TTAGTACATT TATTAAGCGC TAGATTCTGT | 2160 |
| | GCGTTGTTGA TTTACAGACA ATTGTTGTAC GTATTTTAAT AATTCATTAA ATTTATAATC | 2220 |
| | TTTAGGGTGG TATGTTAGAG CGAAAATCAA ATGATTTTCA GCGTCTTTAT ATCTGAATTT | 2280 |
| 25 | AAATATTAAA TCCTCAATAG ATTTGTAAAA TAGGTTTCGA TTAGTTTCAA ACAAGGGTTG | 2340 |
| | TTTTTCCGAA CCGATGGCTG GACTATCTAA TGGATTTTCG CTCAACGCCA CAAAACCTGC | 2400 |
| | CAAATCTTGT AGCAGCAATC TAGCTTTGTC GATATTCGTT TGTGTTTTGT TTTGTAATAA | 2460 |
| 30 | AGGTTTCGACG TCGTTCAAAA TATTATGCGC TTTTGTATTT CTTTCATCAC TGTCGTTAGT | 2520 |
| | GTACAATTGA CTCGACGTAA ACACGTAAAT TAAAGCCTGG ACATATTTAA CATCGGGCGT | 2580 |
| 35 | GTTAGCTTTA TTAGGCCGAT TATCGTCGTC GTCCCAACCC TCGTCGTTAG AAGTTGCTTC | 2640 |
| | CGAAGACGAT TTTGCCATAG CCACACGACG CCTATTAATT GTGTCGGCTA ACACGTCCGC | 2700 |
| | GATCAAATTT GTAGTTGAGC TTTTGAAT TATTTCTGAT TGCGGGCGTT TTTGGGCGGG | 2760 |
| 40 | TTTCAATCTA ACTGTGCCCC ATTTTAATTC AGACAACACG TTAGAAAGCG ATGGTGCAGG | 2820 |
| | CGGTGGTAAC ATTTAGACG GCAAATCTAC TAATGGCGGC GGTGGTGGAG CTGATGATAA | 2880 |
| 45 | ATCTACCATC GGTGGAGGCG CAGGCGGGGC TGCCGGCGGA GGCGGAGGCG GAGGTGGTGG | 2940 |
| | CGGTGATGCA GACGGCGGTT TAGGCTCAAA TTGTCTCTTT CAGGCAACAC AGTCGGCACC | 3000 |
| | TCAACTATTG TACTGGTTTC GGGCGTATGG TGCACTCTCA GTACAATCTG CTCTGATGCC | 3060 |
| 50 | GCATAGTTAA GCCAGCCCCG ACACCCGCCA ACACCCGCTG ACGCGCCCTG ACGGGCTTGT | 3120 |
| | CTGCTCCCGG CATCCGCTTA CAGACAAGCT GTGACCGTCT CCGGGAGCTG CATGTGTCAG | 3180 |

| | | |
|----|--|------|
| | AGGTTTTTCAC CGTCATCACC GAAACGCGCG AGACGAAAGG GCCTCGTGAT ACGCCTATTT | 3240 |
| | TTATAGGTTA ATGTCATGAT AATAATGGTT TCTTAGACGT CAGGTGGCAC TTTTCGGGGA | 3300 |
| 5 | AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC | 3360 |
| | ATGAGACAAT AACCCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT | 3420 |
| 10 | CAACATTTCC GTGTCGCCCT TATTCCTTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT | 3480 |
| | CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT | 3540 |
| | TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT | 3600 |
| 15 | TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTATTGAC | 3660 |
| | GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC | 3720 |
| | TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT | 3780 |
| 20 | GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG | 3840 |
| | AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG | 3900 |
| 25 | GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA | 3960 |
| | ATGGCAACAA CGTTGCGCAA ACTATTAAC TGGCGAACTAC TTA CTCTAGC TTCCCGGCAA | 4020 |
| | CAATTAATAG ACTGGATGGA GCGGATAAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT | 4080 |
| 30 | CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC | 4140 |
| | ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG | 4200 |
| | AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT | 4260 |
| 35 | AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTAAAACTT | 4320 |
| | CATTTTAAAT TAAAAAGGAT CTAGGTGAAG ATCCTTTTGT ATAATCTCAT GACCAAAATC | 4380 |
| 40 | CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT | 4440 |
| | TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA | 4500 |
| | CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTCCGAA GGTAAGTGGC | 4560 |
| 45 | TTCAGCAGAG CGCAGATACC AAATACTGTT CTTCTAGTGT AGCCGTAGTT AGGCCACCAC | 4620 |
| | TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT | 4680 |
| | GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT | 4740 |
| 50 | AAGGCGCAGC GGTGGGCTG AACGGGGGGT TCGTGACAC AGCCCAGCTT GGAGCGAACG | 4800 |
| | ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA | 4860 |

GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG 4920
 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA 4980
 5 CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC 5040
 AACGCGGCCT TTTTACGGT CTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTTCCT 5100
 10 GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT 5160
 CGCCGCAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCATC CTGCACCATC 5220
 GTCTGCTCAT CCATGACCTG ACCATGCAGA GGATGATGCT CGTGACGGTT AACGCCTCGA 5280
 15 ATCAGCAACG GCTTGCCGTT CAGCAGCAGC AGACCATTTT CAATCCGCAC CTCGCGGAAA 5340
 CCGACATCGC AGGCTTCTGC TTCAATCAGC GTGCCGTCGG CGGTGTGCAG TTCAACCACC 5400
 GCACGATAGA GATTCGGGAT TTCGGCGCTC CACAGTTTCG GGTTTTCGAC GTTCAGACGT 5460
 20 AGTGTGACGC GATCGGTATA ACCACCACGC TCATCGATAA TTTCACCGCC GAAAGGCGCG 5520
 GTGCCGCTGG CGACCTGCGT TTCACCCTGC CATAAAGAAA CTGTTACCCG TAGGTAGTCA 5580
 25 CGCAACTCGC CGCACATCTG AACTTCAGCC TCCAGTACAG CGCGGCTGAA ATCATCATT 5640
 AAGCGAGTGG CAACATGGAA ATCGTGATT TGTGTAGTCG GTTTATGCAG CAACGAGACG 5700
 TCACGAAAAA TGCCGCTCAT CCGCCACATA TCCTGATCTT CCAGATAACT GCCGTCACCTC 5760
 30 CAACGCAGCA CCATCACCGC GAGGCGGTTT TCTCCGGCGC GTAAAAATGC GCTCAGGTCA 5820
 AATTCAGACG GCAAACGACT GTCCTGGCCG TAACCGACCC AGCGCCCGTT GCACCACAGA 5880
 35 TGAAACGCCG AGTTAACGCC ATCAAAAATA ATTTCGCTCT GGCCTTCCTG TAGCCAGCTT 5940
 TCATCAACAT TAAATGTGAG CGAGTAACAA CCCGTCGGAT TCTCCGTGGG AACAAACGGC 6000
 GGATTGACCG TAATGGGATA GGTACCGTTG GTGTAGATGG GCGCATCGTA ACCGTGCATC 6060
 40 TGCCAGTTTG AGGGGACGAC GACAGTATCG GCCTCAGGAA GATCGCACTC CAGCCAGCTT 6120
 TCCGGCACCG CTTCTGGTGC CGGAAACCAG GCAAAGCGCC ATTCGCCATT CAGGCTGCGC 6180
 AACTGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT TACGCCAGCT GGCGAAAGGG 6240
 45 GGATGTGCTG CAAGGCGATT AAGTTGGGTA ACGCCAGGGT TTTCCAGTC ACGACGTTGT 6300
 AAAACGACGG GATCTATCAT TTTTAGCAGT GATTCTAATT GCAGCTGCTC TTTGATACAA 6360
 50 CTAATTTTAC GACGACGATG CGAGCTTTTA TTCAACCGAG CGTGCATGTT TGCAATCGTG 6420
 CAAGCGTTAT CAATTTTCA TTATCGTATT GTTGACATC AACAGGCTGG ACACCACGTT 6480

GAACTCGCCG CAGTTTTGCG GCAAGTTGGA CCCGCCGCGC ATCCAATGCA AACTTTCCGA 6540
 CATTCTGTTG CCTACGAACG ATTGATTCTT TGTCCATTGA TCGAAGCGAG TGCCTTCGAC 6600
 TTTTTCGTGT CCAGTGTGGC TT 6622

(2) ANGABEN ZU SEQ ID NO: 43:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 31 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: exon
- (B) LAGE: 1..31

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 43:

CCGGATCCGC CCAGGGCCAC CTAAGGAGCG G 31

(2) ANGABEN ZU SEQ ID NO: 44:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 29 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Einzelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

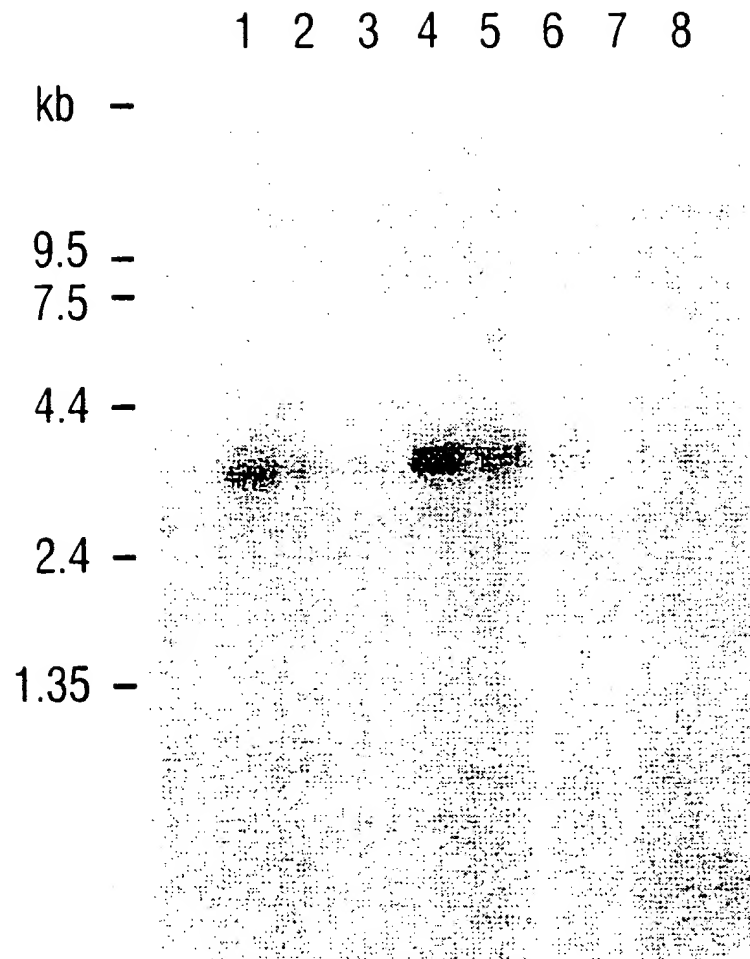
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 44:

CTGAATTCAG GAGCCAGGGC ACAGGCATG 29

Patentansprüche

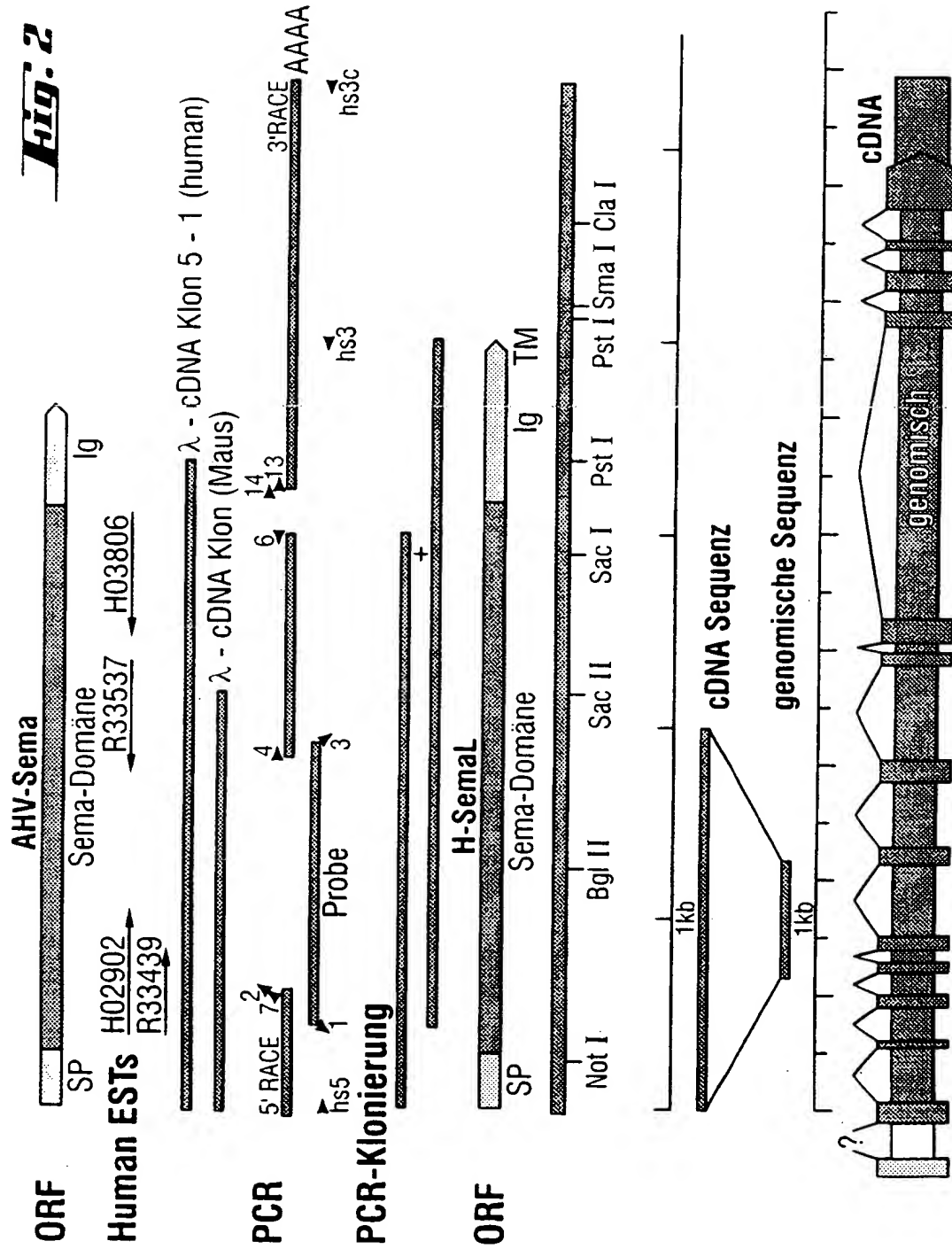
1. Semaphorin enthaltend eine charakteristische Sema-Domäne, dadurch gekennzeichnet, daß das Protein ein N-terminales Signalpeptid und im C-terminalen Bereich eine Immunglobulin-ähnliche Domäne und eine Transmembrandomäne aufweist, wobei das Semaphorin als Semaphorin vom Typ L (SemaL) bezeichnet wird und Derivate des Semaphorins vom Typ L.
2. Semaphorin nach Anspruch 1, wobei das Protein (humanes Semaphorin vom Typ L (H-SemaL)) die Aminosäuresequenz SEQ ID NO. 3 hat.

3. Semaphorin nach einem oder mehreren der Ansprüche 1 und 2, wobei das Protein im Bereich der Sema-Domäne eine Aminosäureidentität von mindestens 40 % im Bezug auf die Sema-Domäne von H-SemaL aufweist.
- 5 4. Semaphorin nach einem oder mehreren der Ansprüche 1 bis 2, wobei das Protein die partielle Aminosäuresequenz SEQ ID NO. 4 enthält (murines Semaphorin (M-SemaL)).
5. Nukleinsäure enthaltend eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L gemäß einem oder mehreren der Ansprüche 1 bis 4 kodiert sowie Derivate derselben.
- 10 6. Nukleinsäure nach Anspruch 5, wobei besagte Nukleinsäure-Sequenz ein Semaphorin L Gen ist.
7. Nukleinsäure nach einem oder mehreren der Ansprüche 5 und 6, wobei besagte Nukleinsäure-Sequenz das Gen von H-SemaL enthält.
- 15 8. Nukleinsäure nach Anspruch 5, wobei besagte Nukleinsäure-Sequenz die cDNA eines Semaphorins vom Typ L enthält.
9. Nukleinsäure nach Anspruch 8, wobei die cDNA die cDNA von H-SemaL ist.
- 20 10. Nukleinsäure nach Anspruch 8, wobei die cDNA die cDNA von M-SemaL ist.
11. Verfahren zur Herstellung eines Semaphorins vom Typ L gemäß einem oder mehreren der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L oder ein Derivat derselben kodiert, in einen Expressionsvektor kloniert und exprimiert wird.
- 25 12. Verfahren gemäß Anspruch 11, wobei für die Expression eine eukaryotische Zelle verwendet wird.
13. Verwendung eines Semaphorins vom Typ L oder eines Derivats desselben oder einer Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert oder eines Derivats derselben zur Herstellung eines Arzneimittels, das zur Behandlung oder Prävention von immunologischen Erkrankungen verwendet werden kann.
- 30 14. Verwendung einer Nukleinsäure-Sequenz oder eines Derivats derselben nach Anspruch 13 in der Gentherapie.
15. Verwendung eines Semaphorins vom Typ L oder einer Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert in einem Verfahren zur Identifizierung von immunmodulierenden Wirkstoffen.
- 35 16. Verfahren zur Identifizierung von immunmodulierenden Wirkstoffen, dadurch gekennzeichnet, daß ein Semaphorin vom Typ L unter definierten Bedingungen mit einem zu untersuchenden Wirkstoff inkubiert wird, parallel ein zweiter Ansatz ohne den zu untersuchenden Wirkstoff, aber unter ansonsten gleichen Bedingungen durchgeführt wird und dann die inhibierende bzw. aktivierende Wirkung des zu untersuchenden Wirkstoffs bestimmt wird.
- 40 17. Verfahren zur Identifizierung von immunmodulierenden Wirkstoffen, dadurch gekennzeichnet, daß eine Nukleinsäure-Sequenz, die für ein Semaphorin vom Typ L kodiert unter definierten Bedingungen und in Gegenwart eines zu untersuchenden Wirkstoffs exprimiert wird und das Ausmaß der Expression bestimmt wird.
- 45 18. Verfahren zur Herstellung einer Nukleinsäure, die für ein Semaphorin vom Typ L kodiert, wobei diese Nukleinsäure mit Hilfe der Polymerase Ketten Reaktion unter Verwendung von spezifischen Primern amplifiziert wird.
19. Semaphorin Antikörper, dadurch gekennzeichnet, daß er entweder das Epitop von H-SemaL, das den Aminosäuren 179-378 in SEQ ID NO: 4 entspricht, erkennt oder daß er das Epitop von H-SemaL, das den Aminosäuren 480-666 in SEQ ID NO: 4 entspricht, erkennt.
- 50 20. Verfahren zur Herstellung eines Semaphorin Antikörpers nach Anspruch 19, wobei die Epitope als Fusionsproteine mit einem Epitop-Tag exprimiert und über dieses Epitop-Tag aufgereinigt werden und die aufgereinigten Fusionsproteine zur Immunisierung verwendet werden.
- 55

Fig. 1

Beta-Actin
Kontrolle



Fig. 2

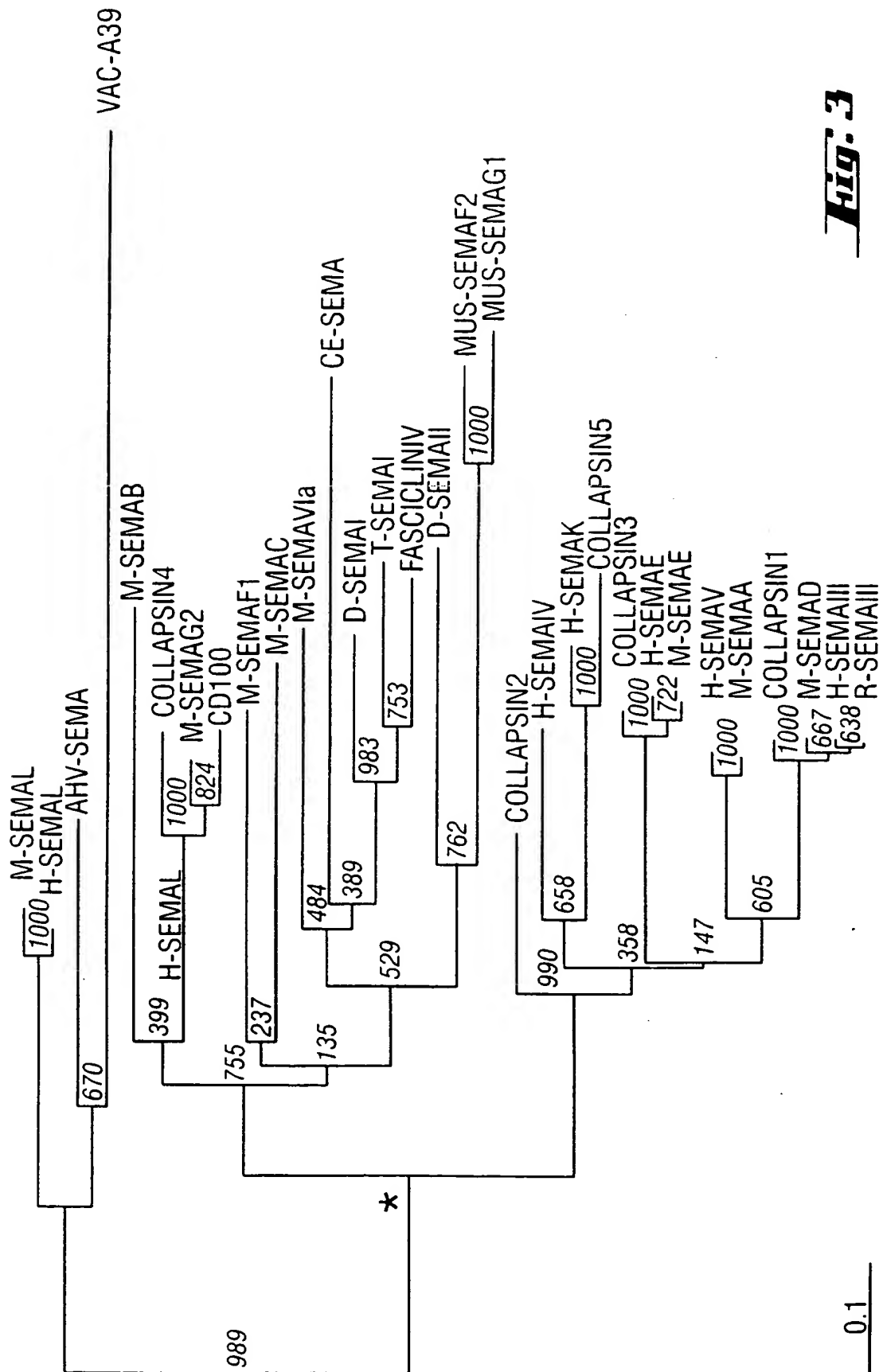


Fig. 3

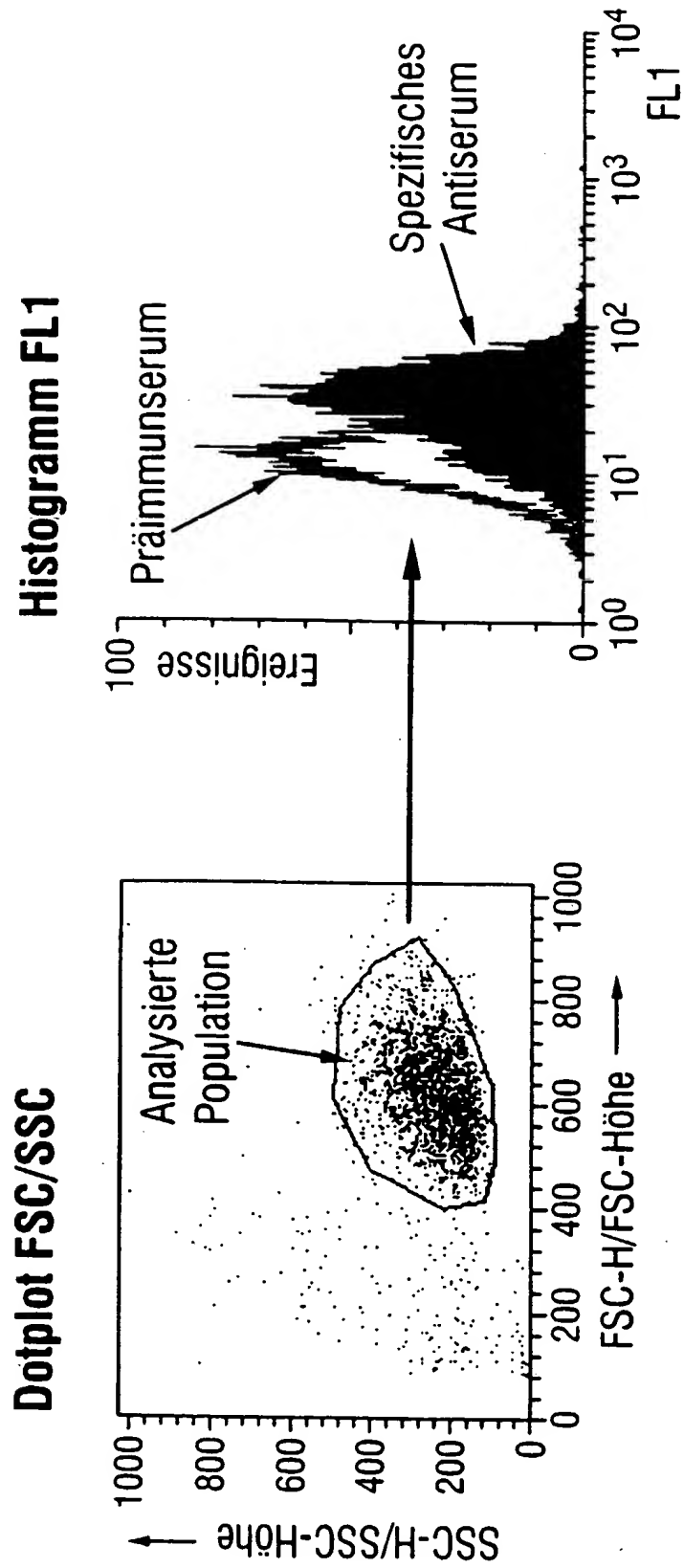
Fig. 4a (U937)

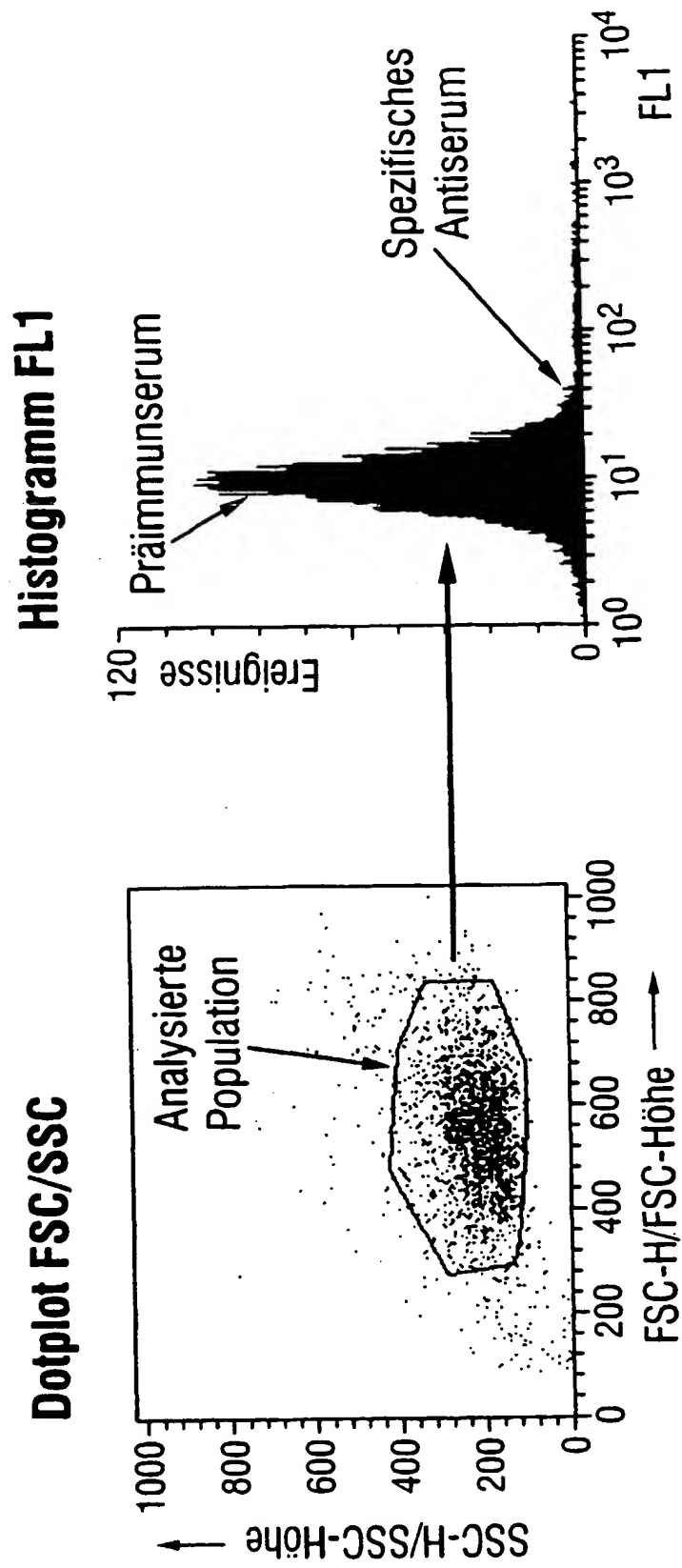
Fig. 4b (THP-1)

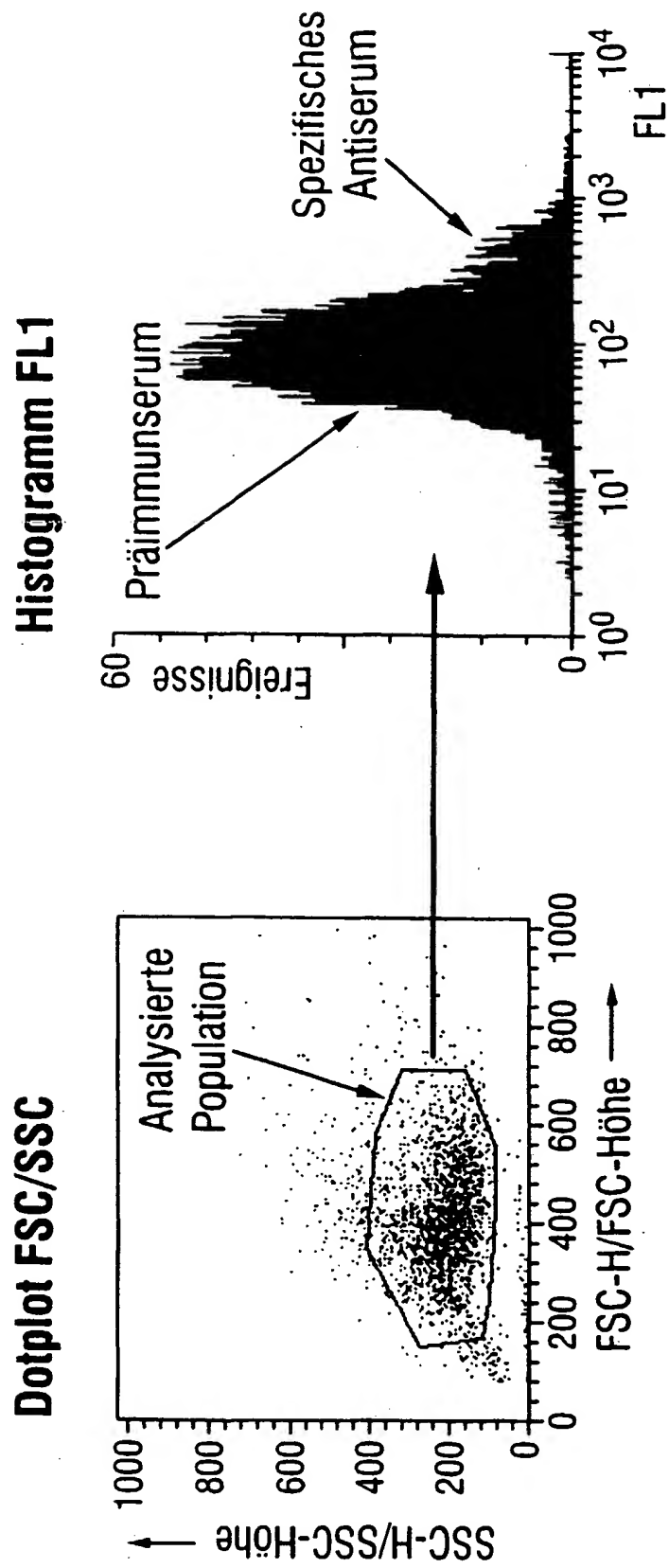
Fig. 4-c (K562)

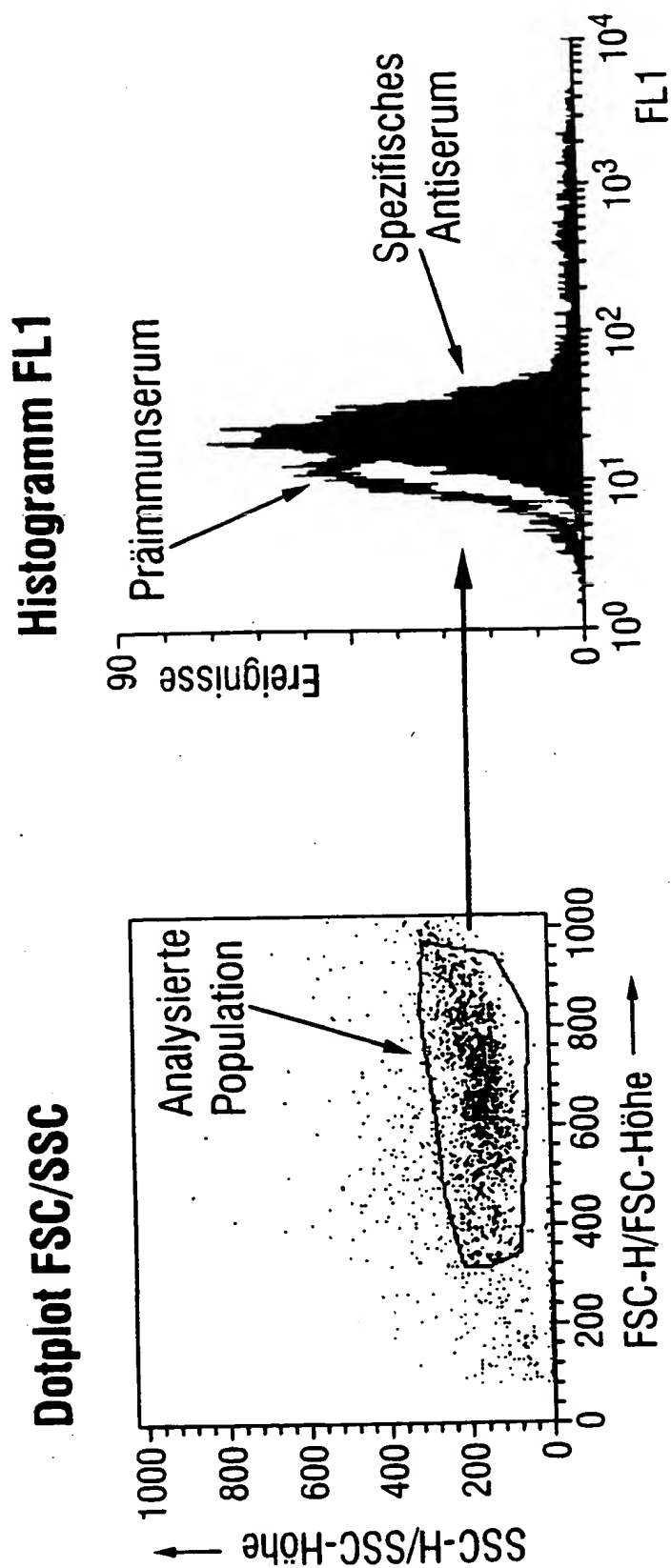
Fig. 4d (L-428)

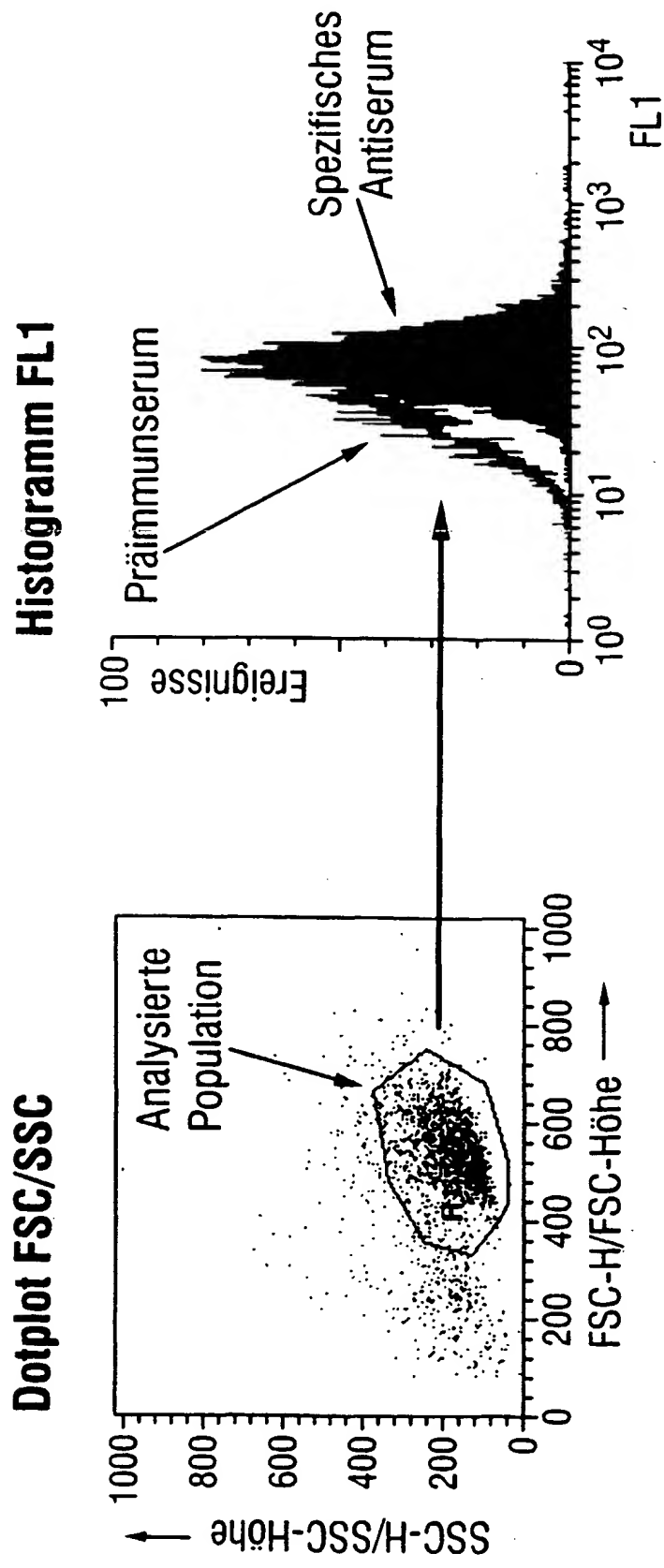
Fig. 4e (Jurkat)

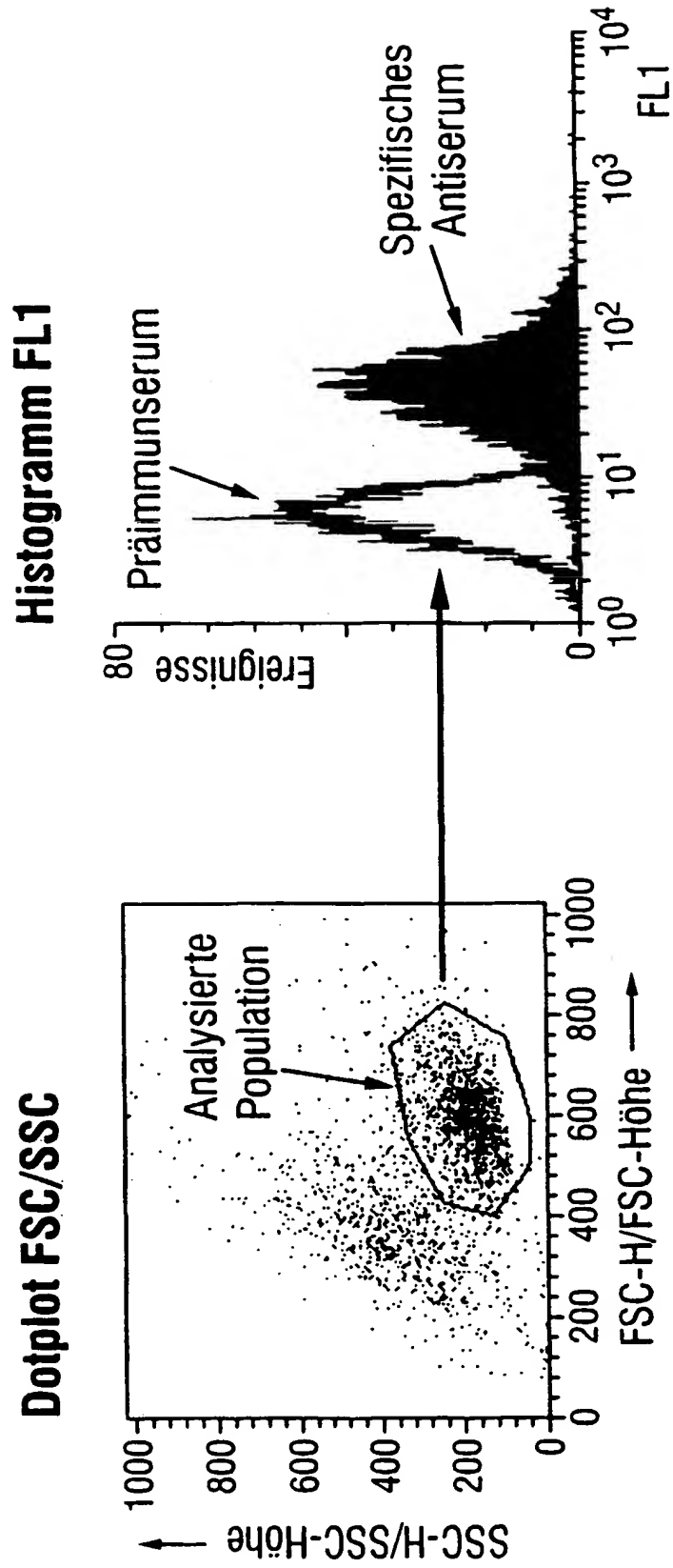
Fig. 4f (Daudi)

Fig. 4g (LCL EBV-Transformierte B-Zellen)

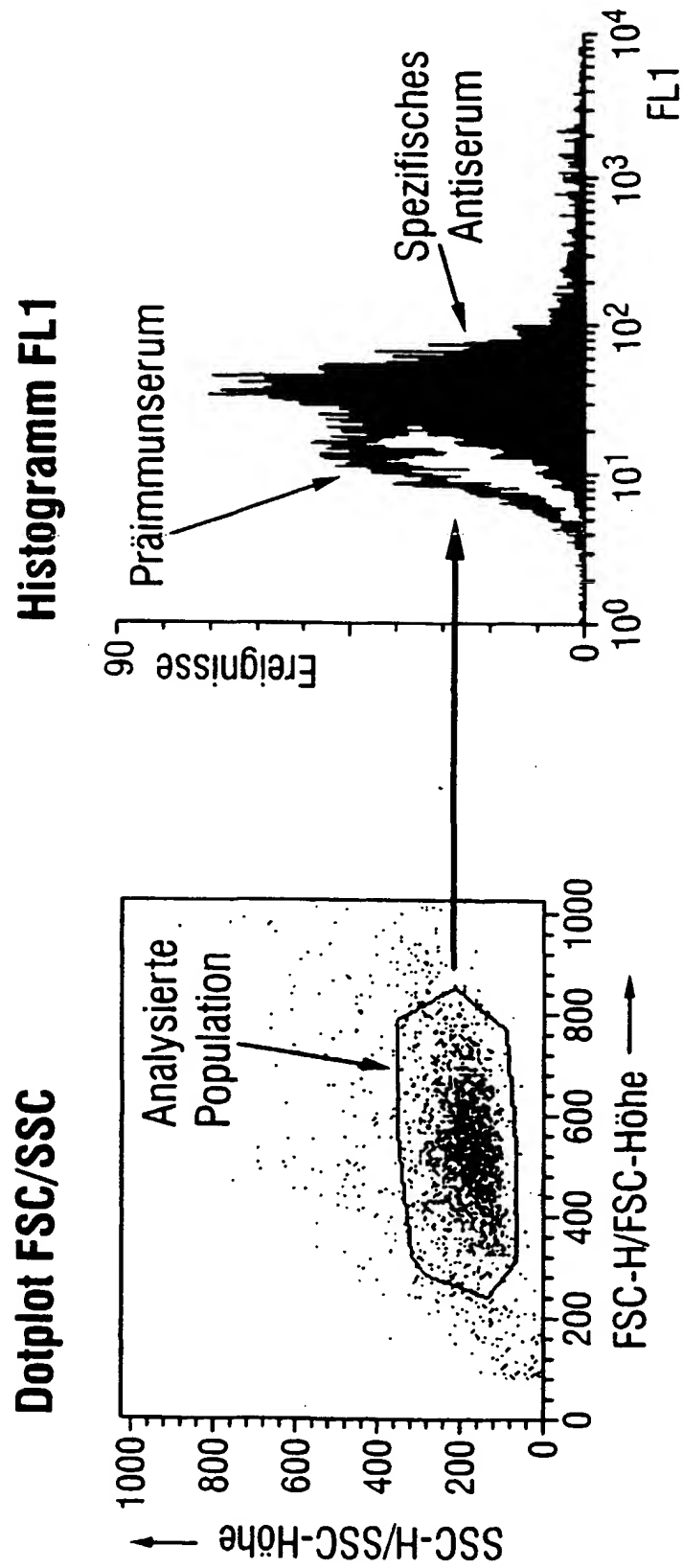


Fig. 4h (Jiyoye)

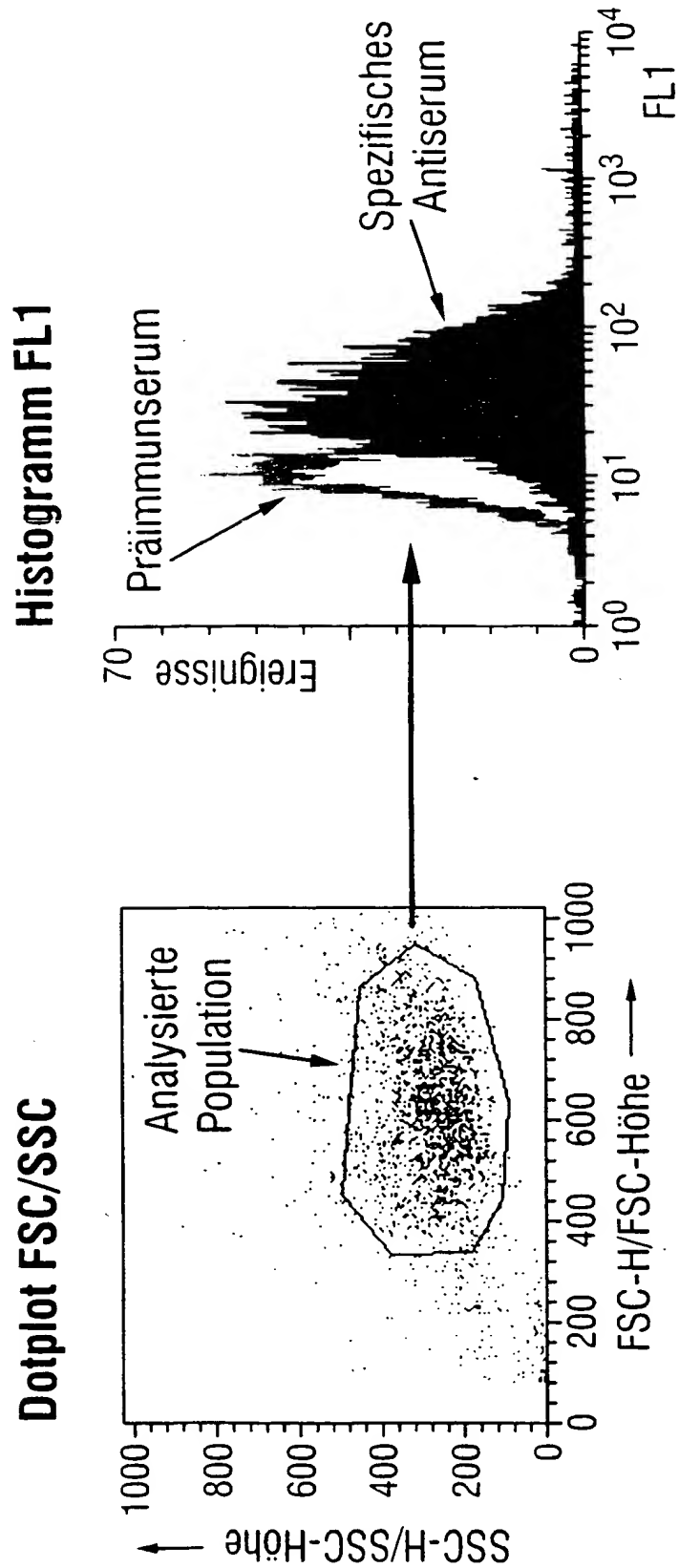


Fig. 4i (CBL-Mix57)

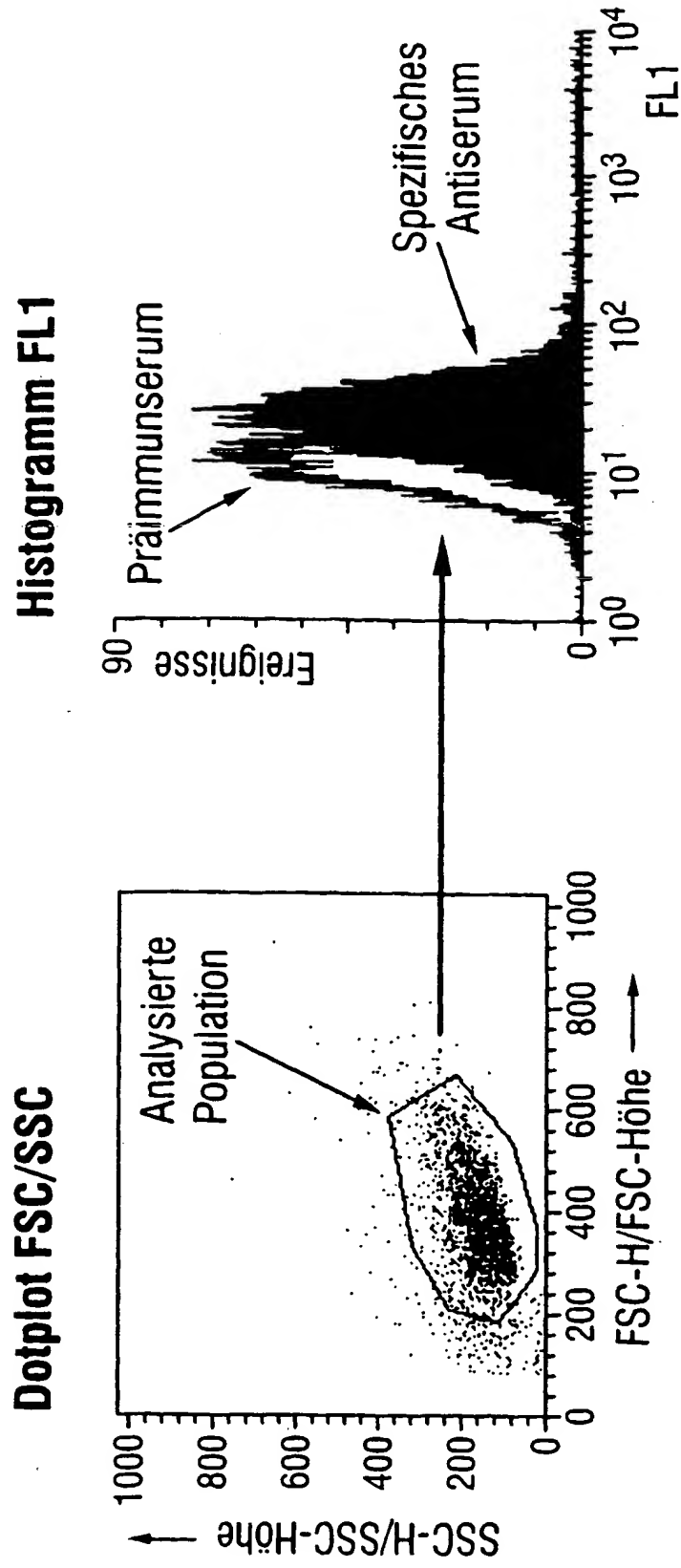
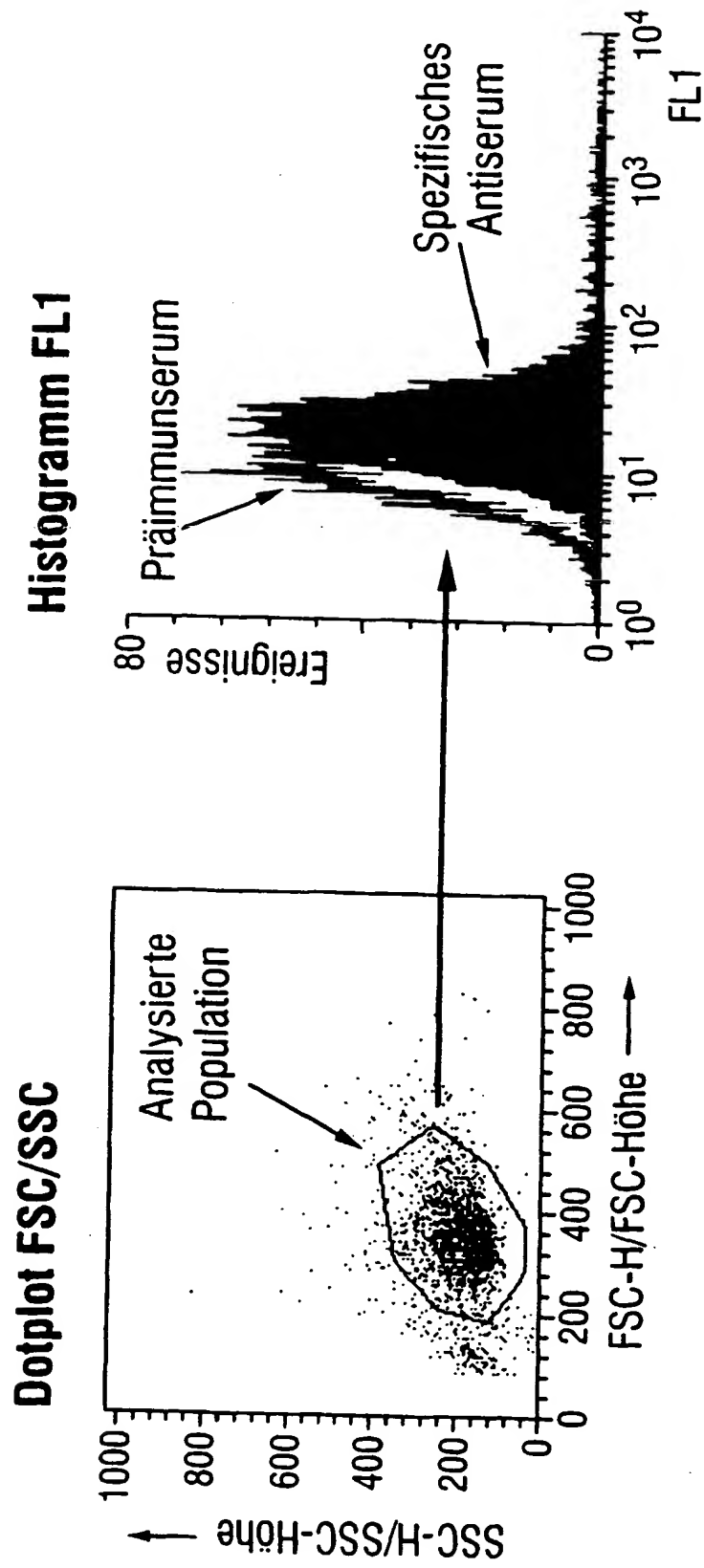
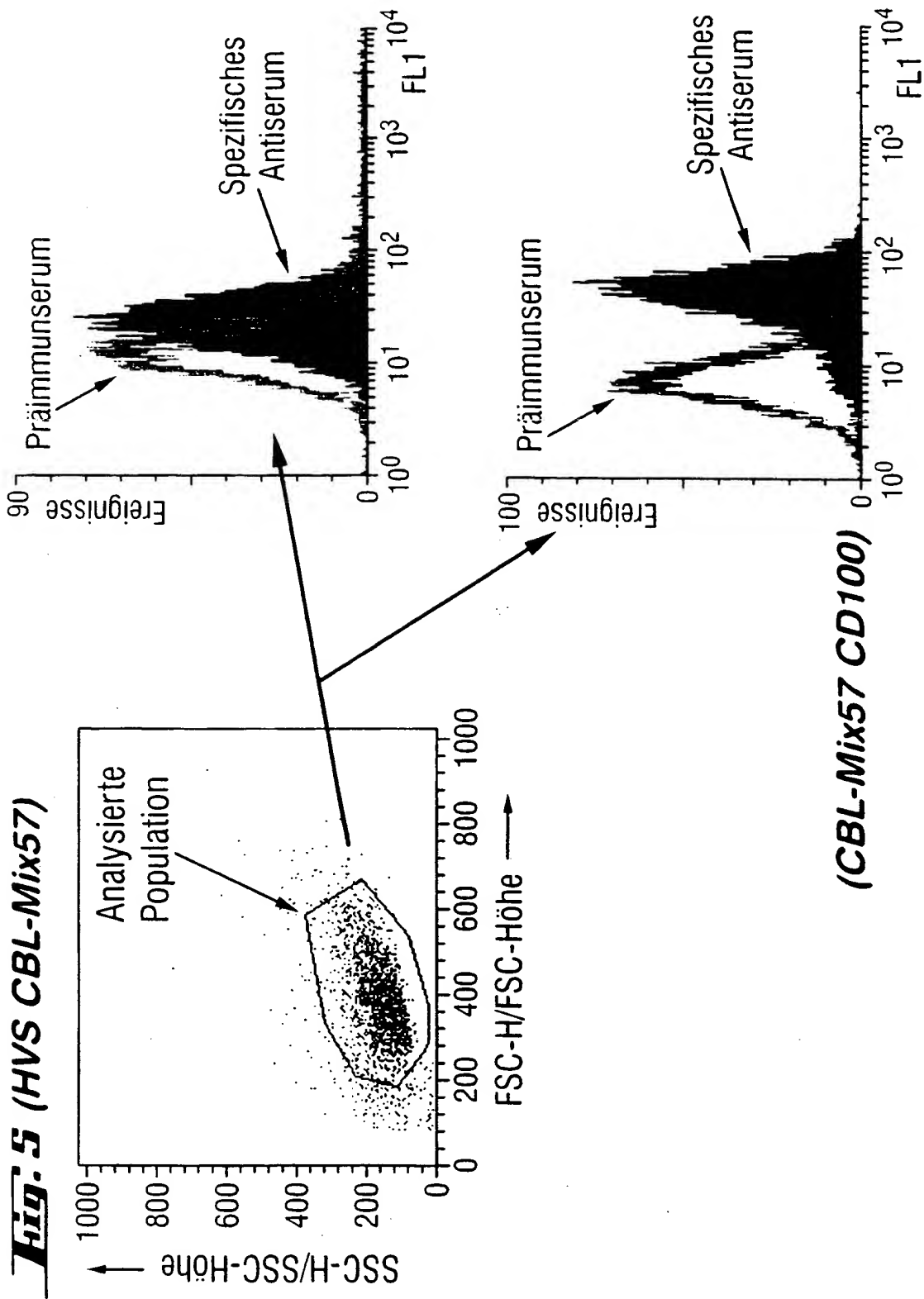


Fig. 4k (CBL-Mix59)





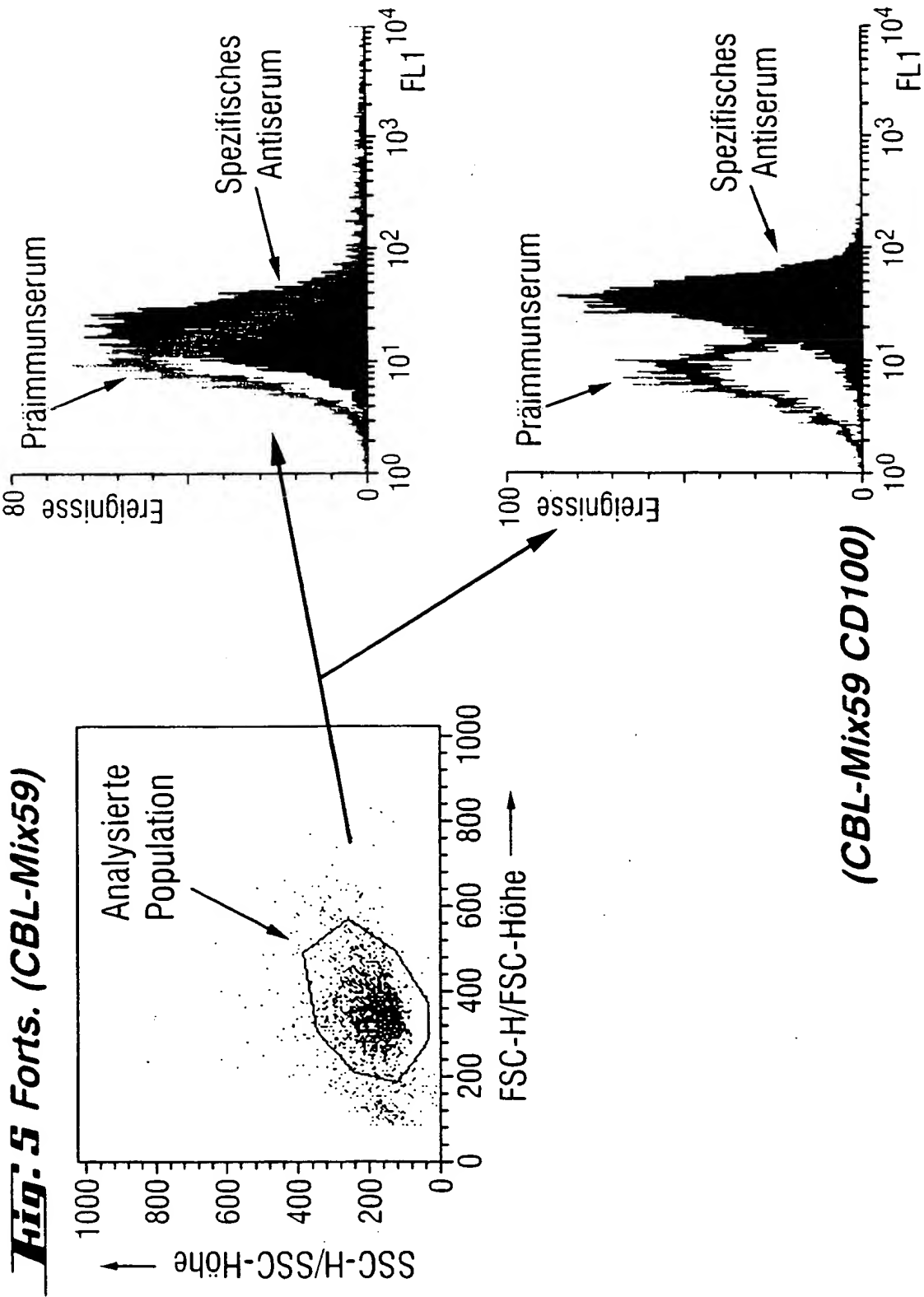


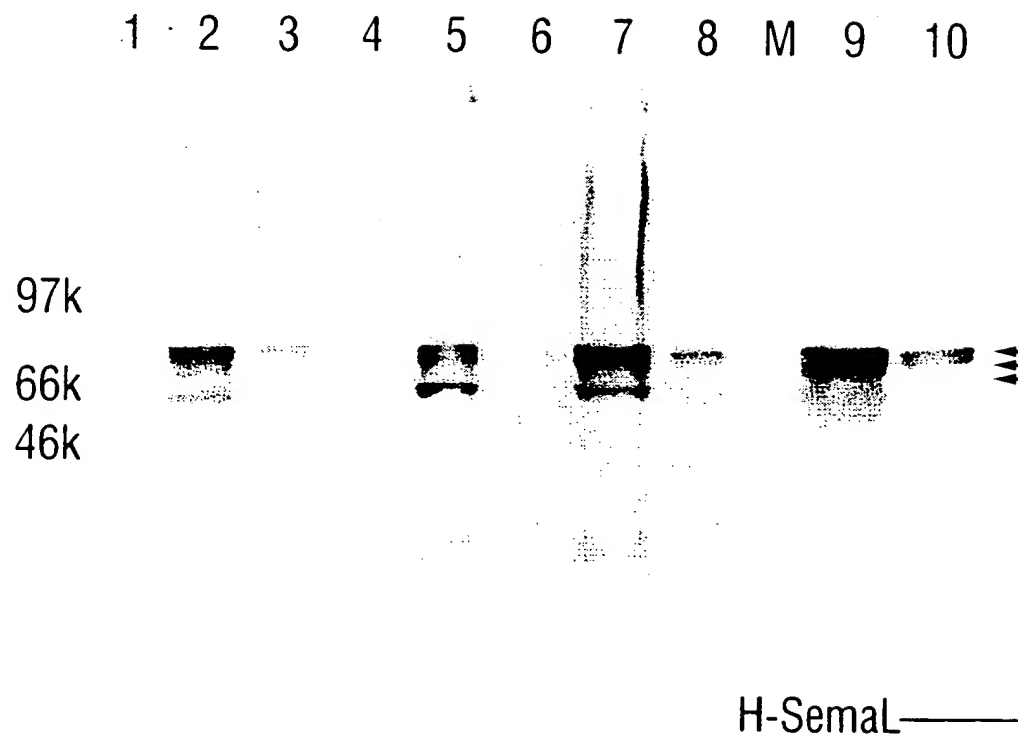
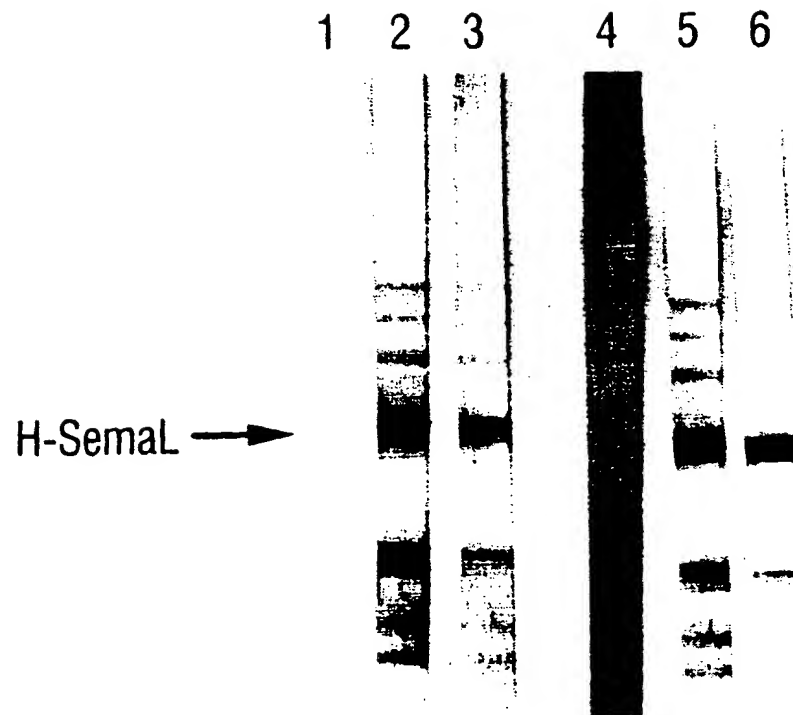
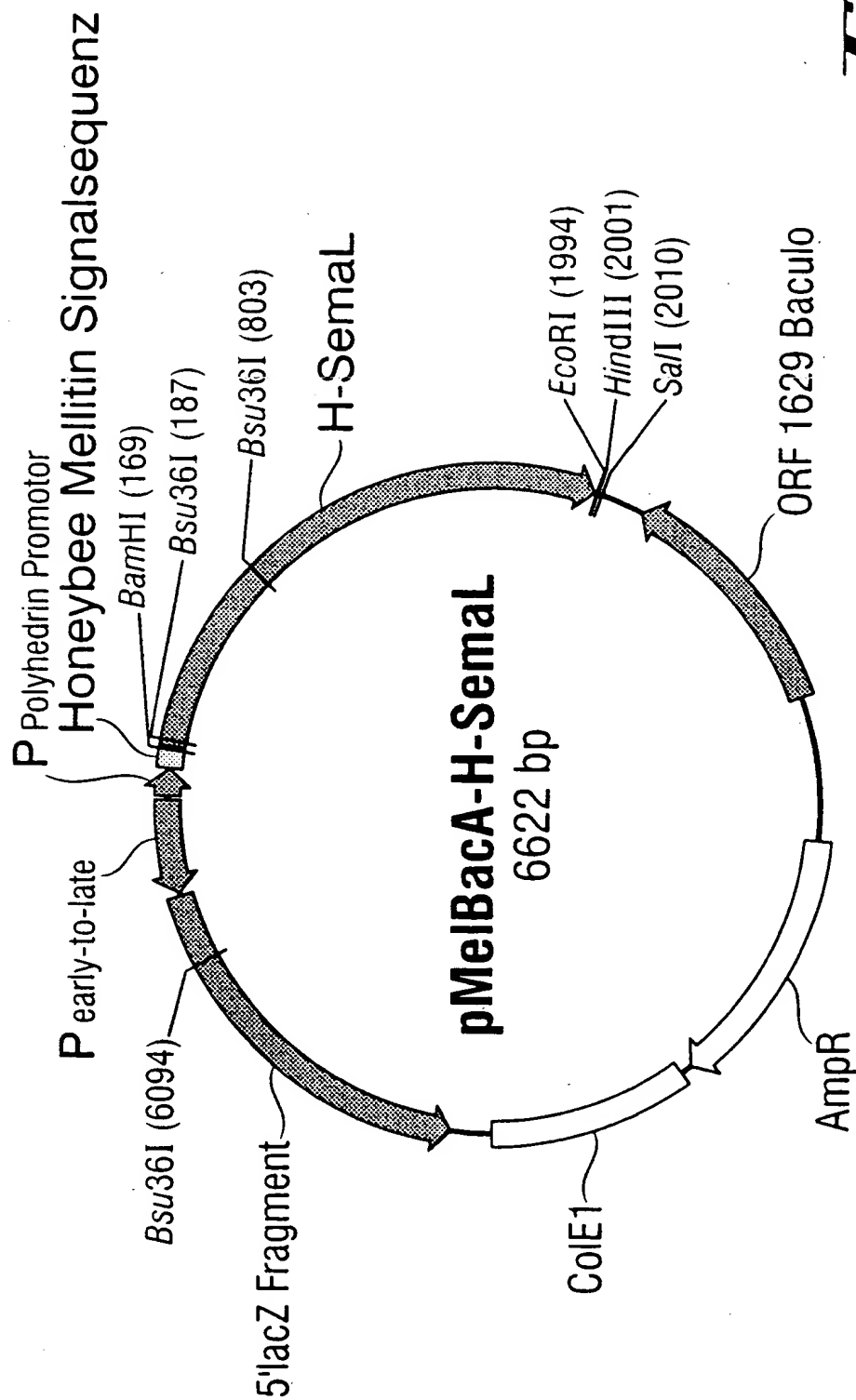
Fig. 6

Fig. 1

**Fig. 8**

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 892 047 A3

(12)

EUROPÄISCHE PATENTANMELDUNG

(88) Veröffentlichungstag A3:
08.03.2000 Patentblatt 2000/10

(51) Int. Cl.⁷: **C12N 15/12**, C07K 14/47,
A61K 48/00, G01N 33/50

(43) Veröffentlichungstag A2:
20.01.1999 Patentblatt 1999/03

(21) Anmeldenummer: **98112470.4**

(22) Anmeldetag: 06.07.1998

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: 09.07.1997 DE 19729211
11.02.1998 DE 19805371

(71) Anmelder:
Hoechst Marion Roussel Deutschland GmbH
65929 Frankfurt am Main (DE)

(72) Erfinder:
• **Fleckenstein, Bernhard, Prof. Dr.**
91369 Wiesenthau (DE)
• **Ensser, Armin, Dr.**
90419 Nürnberg (DE)

(54) **Humanes und murines Semaphorin L**

(57) **Humanes Semaphorin L (H-SemaL) und korrespondierende Semaphorine in anderen Spezies.**

Gegenstand der Erfindung sind neuen Semaphorine, die sich durch eine besondere Domänenstruktur auszeichnen und deren Derivate, Nukleinsäuren (DNA, RNA, cDNA), die für diese Semaphorine kodieren und deren Derivate sowie die Verwendung derselben.

Gegenstand der vorliegenden Erfindung sind Semaphorine mit einer neuen, bisher nicht bekannten und nicht zu erwartenden Domänenstruktur, denen eine biochemische Funktion im Immunsystem zukommt (immunmodulierende Semaphorine). Die erfindungsgemäßen Semaphorine werden als Semaphorine vom Typ L (SemaL) bezeichnet. Sie enthalten ein N-terminales Signalpeptid, eine charakteristische Sema-Domäne und im C-terminalen Bereich des Proteins eine Immunglobulin-ähnliche Domäne und eine hydrophobe Domäne, die eine potentielle Transmembrandomäne darstellt.

EP 0 892 047 A3



Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung
EP 98 11 2470

| EINSCHLÄGIGE DOKUMENTE | | | |
|---|---|--|---|
| Kategorie | Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile | Betrifft Anspruch | KLASSIFIKATION DER ANMELDUNG (Int.Cl.6) |
| D,X | ENSSER A ET AL: "Alcephaline herpesvirus type 1 has a semaphorin-like gene" JOURNAL OF GENERAL VIROLOGY., Bd. 76, 1995, Seiten 1063-1067, XP002113888 SOCIETY FOR GENERAL MICROBIOLOGY, READING., GB ISSN: 0022-1317 * Abbildung 2 * | 1,3,5,6, 8,11 | C12N15/12 C07K14/47 A61K48/00 G01N33/50 |
| D,A | DATABASE EMBL 'Online! Accession Number H03806, 22. Juni 1995 (1995-06-22) HILLIER L ET AL: "yj39f01.r1 Homo sapiens cDNA clone 151129 5' similar to SP:A49069 A49069 COLLAPSIN - ;" XP002125533 * 94% identität (420 BP) zwischen SEQ ID 1 und H03806 * | 5 | |
| A | DATABASE EMBL 'Online! accession number AA260340, 19. März 1997 (1997-03-19) MARRA M ET AL: "va91b02.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 746763 5' similar to TR:G1000717 G1000717 SIMILAR TO GENBANK ACCESSION NUMBER L26081. ;" XP002125534 * 99.7% identität (347 BP) zwischen SEQ ID 1 und H03806 * | 4 | RECHERCHIERTE SACHGEBIETE (Int.Cl.6) C12N C07K A61K G01N |
| Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt | | | |
| Recherchenort | Abschlußdatum: der Recherche | Prüfer | |
| DEN HAAG | 10. Januar 2000 | Lejeune, R | |
| KATEGORIE DER GENANNTEN DOKUMENTE | | | |
| X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : mündliche Offenbarung P : Zwischenliteratur | | T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentedokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument | |



Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung
EP 98 11 2470

| EINSCHLÄGIGE DOKUMENTE | | | |
|--|--|--|---|
| Kategorie | Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile | Betrifft Anspruch | KLASSIFIKATION DER ANMELDUNG (Int.Cl.6) |
| T | <p>YAMADA A ET AL: "Molecular cloning of a glycosylphosphatidylinositol-anchored molecule CDw108"</p> <p>JOURNAL OF IMMUNOLOGY.,</p> <p>Bd. 162, 1. April 1999 (1999-04-01),</p> <p>Seiten 4094-4100, XP002123609</p> <p>THE WILLIAMS AND WILKINS CO. BALTIMORE., US</p> <p>ISSN: 0022-1767</p> <p>* 99.8% Identität zwischen SEQ ID 1 und CDW108 (Abbildung 1) *</p> <p>* Seite 4094, Spalte 2, Absatz 2 *</p> <p>---</p> | <p>1-3,5-9,</p> <p>11,12,</p> <p>18-20</p> | |
| T | <p>LANGE C ET AL: "New eukaryotic semaphorins with close homology to semaphorins of viruses"</p> <p>GENOMICS,</p> <p>Bd. 51, 1. August 1998 (1998-08-01),</p> <p>Seiten 340-350, XP002113887</p> <p>SAN DIEGO., US</p> <p>ISSN: 0888-7543</p> <p>* 99.8% Identität zwischen SEQ ID 1 und H-Sema-L (Abbildung 3) *</p> <p>-& DATABASE EMBL 'Online!</p> <p>Accession Number AF030699,</p> <p>8. September 1998 (1998-09-08)</p> <p>LANGE C ET AL: "Mus musculus semaphorin L (SemaL) mRNA, partial cds."</p> <p>XP002125537</p> <p>* M-SemaL ist identisch zu SEQ ID 1 *</p> <p>---</p> <p>-/--</p> | <p>1-12,18</p> | <p>RECHERCHIERTE SACHGEBIETE (Int.Cl.6)</p> |
| Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt | | | |
| Recherchenort | | Abschlußdatum der Recherche | Prüfer |
| DEN HAAG | | 10. Januar 2000 | Lejeune, R |
| <p>KATEGORIE DER GENANNTEN DOKUMENTE</p> <p>X : von besonderer Bedeutung allein betrachtet</p> <p>Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie</p> <p>A : technologischer Hintergrund</p> <p>O : nichtschriftliche Offenbarung</p> <p>P : Zwischenliteratur</p> <p>T : der Erfindung zugrunde liegende Theorien oder Grundsätze</p> <p>E : älteres Patentedokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist</p> <p>D : in der Anmeldung angeführtes Dokument</p> <p>L : aus anderen Gründen angeführtes Dokument</p> <p>& : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument</p> | | | |



Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung
EP 98 11 2470

| EINSCHLÄGIGE DOKUMENTE | | | |
|---|--|---|---|
| Kategorie | Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile | Betrifft Anspruch | KLASSIFIKATION DER ANMELDUNG (Int.Cl.6) |
| T | XU X ET AL: "Human semaphorin K1 is glycosylphosphatidylinositol-linked and defines a new subfamily of viral-related semaphorins" JOURNAL OF BIOLOGICAL CHEMISTRY., Bd. 273, Nr. 35, 28. August 1998 (1998-08-28), Seiten 22428-22434, XP002113886 AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD., US ISSN: 0021-9258 * 99.1% Identität zwischen SEQ ID 1 und Sema K1 (Abbildung 1) * | 1-3,5-9, 11,12,18 | |
| E | WO 99 45114 A (ZYMOGENETICS INC) 10. September 1999 (1999-09-10) * 99.8% Identität zwischen SEQ ID 1 und SEQ ID 1 von W09945114) * * Beispiel 4 * | 1-3,5-9, 11,12, 18-20 | |
| E | WO 99 38885 A (SMITHKLINE BEECHAM PLC) 5. August 1999 (1999-08-05) * 99.9% Identität zwischen SEQ ID 1 und SEQ ID 1 von W09938885 * | 1-3,5-9, 11,12,18 | RECHERCHIERTE SACHGEBIETE (Int.Cl.6) |
| E | WO 99 58676 A (SPRIGGS MELANIE K ;IMMUNEX CORP (US)) 18. November 1999 (1999-11-18) * 99.95% Identität zwischen SEQ ID 1 und SEQ ID 1 von W09958676 * | 1-3,5-9, 11,12,18 | |
| -/-- | | | |
| Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt | | | |
| Recherchenort DEN HAAG | | Abschlußdatum der Recherche 10. Januar 2000 | Prüfer Lejeune, R |
| KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument | | | |



Europäisches
Patentamt

EUROPÄISCHER RECHERCHENBERICHT

Nummer der Anmeldung
EP 98 11 2470

| EINSCHLÄGIGE DOKUMENTE | | | |
|--|--|---|---|
| Kategorie | Kennzeichnung des Dokuments mit Angabe, soweit erforderlich, der maßgeblichen Teile | Betrifft Anspruch | KLASSIFIKATION DER ANMELDUNG (Int.Cl.6) |
| T | DATABASE EMBL 'Online! Accession Number AF176670, 7. September 1999 (1999-09-07) MINE T ET AL: "Mus musculus GPI-anchored membrane protein CDw108 precursor (CDw108) mRNA, complete cds." XP002125535 * 99.8% identität zwischen SEQ ID 2 und AF176670 * | 1,3-6,8, 10-12,18 | |
| T | DATABASE EMBL 'Online! Accession number ab017532, 15. März 1999 (1999-03-15) TAKAHASHI H ET AL: "Mus musculus mRNA for msemK1p, complete cds." XP002125536 * 99.8% identität zwischen SEQ ID 2 und AB017532 * | 1,3-6,8, 10-12,18 | |
| | | | RECHERCHIERTE SACHGEBIETE (Int.Cl.6) |
| Der vorliegende Recherchenbericht wurde für alle Patentansprüche erstellt | | | |
| Recherchenort DEN HAAG | | Abschlußdatum der Recherche 10. Januar 2000 | Prüfer Lejeune, R |
| KATEGORIE DER GENANNTEN DOKUMENTE X : von besonderer Bedeutung allein betrachtet Y : von besonderer Bedeutung in Verbindung mit einer anderen Veröffentlichung derselben Kategorie A : technologischer Hintergrund O : nichtschriftliche Offenbarung P : Zwischenliteratur | | T : der Erfindung zugrunde liegende Theorien oder Grundsätze E : älteres Patentdokument, das jedoch erst am oder nach dem Anmeldedatum veröffentlicht worden ist D : in der Anmeldung angeführtes Dokument L : aus anderen Gründen angeführtes Dokument & : Mitglied der gleichen Patentfamilie, übereinstimmendes Dokument | |